Vol. 84, No. 1 Founded in 1887 by G. Stanley Hall March 1971

K.5.
Docket & issue to the Librarian.
R Das
13/7/71

THE AMERICAN JOURNAL OF PSYCHOLOGY

Bureau of Ednl. & Psyl. Research
(S. C. E. R. T.)
Date 3·3·81
Acc J 652

Edited by

Lloyd G. Humphreys **M. E. Bitterman**
Charles W. Eriksen **Benton J. Underwood**

With the cooperation of Karl M. Dallenbach, William Epstein, Rand Evans, Robert Fox, Wendell Garner, J. P. Guilford, Harry Helson, E. R. Hilgard, Julian Hochberg, Francis W. Irwin, F. Nowell Jones, John Krauskopf, G. L. Kreezer, W. C. H. Prentice, Allan Wagner

---

JACOB BECK. Surface Lightness and Cues for the Illumination 1
ERNEST N. DAMIANOPOULOS. Stimulus and Response Availability in Associative Learning 12
THEODORE W. LOREI and JACOB COHEN. A Comparison of Two Methods of Value Measurement 23
DONALD G. MACKAY. Stress Pre-Entry in Motor Systems 35
WILLIAM S. CAIN and JOSEPH C. STEVENS. Effort in Sustained and Phasic Handgrip Contractions 52

continued on back cover

---

vol. 84
1971

The American Journal of Psychology was founded in the interest of general experimental psychology and devotes its pages to reports of original research, to shorter notes and discussions, and to book reviews pertinent to the area. Manuscripts may be submitted to any one of the following:

LLOYD G. HUMPHREYS
Psychology Building
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820

M. E. BITTERMAN
Department of Psychology
University of South Florida
Tampa, Florida 33620

CHARLES W. ERIKSEN
Psychology Building
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820

BENTON J. UNDERWOOD
Department of Psychology
Northwestern University
Evanston, Illinois 60201

All books intended for review should be sent to the editor of that section:

JULIAN HOCHBERG
Department of Psychology
Columbia University
New York, New York 10027

Contributors are asked to follow the style described in the *Publication Manual of the American Psychological Association*. They should keep abstract and title quite brief, use footnotes only when essential, avoid all but the most commonly recognized or proper abbreviations (2 ft, 4 deg, ANOVA, BD-IX rate; *but* subjects, observers, experimenter, paired-associate learning, afterimage), and use the tables and figures to supplement rather than repeat the text. Authors of articles, notes, and discussions are given 100 complimentary copies without covers; covers and additional copies may be ordered at the time galley is returned.

The subscription price is $10.00 a year (foreign $10.50); single copies cost $2.75. Back numbers and volumes may also be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which address all inquiries of a business nature (subscriptions, advertisements, changes of address) should be directed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

© 1971 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

23 JUL 1971

Vol. 84, No. 1 *continued from front cover* March 1971

WILLIAM EPSTEIN. Adaptation to Uniocular Image Magnification after Varying Preadaptation Activities 66

ROBERT TEGHTSOONIAN and MARTHA TEGHTSOONIAN. Discontinuities in Recognition Learning Revealed by Critical-Trial Analysis 75

ARTHUR S. REBER and RICHARD B. MILLWARD. Event Tracking in Probability Learning 85

JAMES R. ERICKSON. Problem Shifts and Hypothesis Behavior in Concept Identification 100

THOMAS W. TURNAGE and JENNY I. STEINMETZ. Unit-Sequence Interference and Short-term Recall 112

EDWIN MARTIN and STEPHEN T. CAREY. Retroaction, Recovery, and Stimulus Meaningfulness in the A-B, A-Br Paradigm 123

DEVENDRA SINGH. Effect of Level of Manifest Anxiety and Type of Pretraining on Timing Behavior 134

Book Reviews 140

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS
Urbana, Illinois 61801

Diary No. 302
Date 13.7.71
File No. Lib
Bureau Psy. Research

SURFACE LIGHTNESS AND CUES FOR THE ILLUMINATION

Jacob Beck
University of Oregon

Observers who judged the lightness of two equally illuminated targets, one of which cast a shadow on a background and one of which did not, perceived the target that cast a shadow (and thus seemed illuminated by a spotlight) as darker than the other target. The results confirm earlier findings that the perception of surface lightness may be affected by cues for the conditions of illumination as well as by the inhibitory interactions underlying lightness contrast. The hypothesis is that such cues affect the perception of lightness by creating the impression of a 'special' illumination.

Numerous experiments have attempted to show that cues for illumination may affect the perception of surface lightness (see, for example, MacLeod, 1932; Kardos, 1934; Gelb, 1929; and Beck, 1965). A basic paradigm was introduced by Gelb and has been used with modifications by both Kardos and MacLeod. In a room illuminated by a weak ceiling lamp Gelb focused the concealed beam of a projection lantern on a revolving black disk (1929). The disk was seen as white and dimly illuminated. When a small piece of white paper was placed to intercept the beam from the lantern, the percept changed. The disk was now seen as black, the paper white, and both as strongly illuminated. The change in the appearance of the disk from a dimly illuminated white to a strongly illuminated black has been ascribed to the white paper's revealing that the disk was strongly illuminated (Woodworth and Schlosberg, 1954, p. 442).

Kardos later conducted a parallel experiment using a concealed shadow (1934). Kardos cast a shadow on a white disk so that there was no penumbra to indicate the shadow. The disk was seen as unshadowed and black. But when the shadow caster was moved slightly to the side so that the edge of the penumbra became visible, the disk was seen as white with a shadow over it. Similarly, MacLeod has shown that casting a shadow on a background, to suggest that a disk standing in front of the background was in shadow, caused the disk to appear lighter (1932). A light gray disk that stood in a concealed shadow appeared black when no shadow was present on the background. When a visible shadow was cast on the background, the disk was seen as a lighter gray in shadow.

These experiments cannot be interpreted unambiguously. The changes in the experimental conditions that altered perceived lightness also altered the luminances in the field. Changes in perceived lightness may be the result, therefore, of contrast effects rather than of the information about illumination. In the Gelb experiment, for example, the introduction of a white paper not only reveals the presence of a hidden light source but also introduces an area of much higher luminance, which, on the basis of contrast, would be expected to darken the disk. Stewart found that the perceived lightness of a disk was a function of the size and position of the small white stimulus introduced into the field, the disk becoming darker as the stimulus was either made larger or moved toward the center of the disk (1959). Newson reported similar results (1958). This indicates that the darkening of the disk in the Gelb experiment is at least in part due to contrast.

The lightening of the targets in the Kardos and MacLeod experiments may also be ascribed to contrast. Making the penumbra visible in Kardos's study changed the contour separating the disk from the background from sharp to graded; and as O'Brien (1958) and MacLeod (1947) have shown, a graded contour has the effect of decreasing contrast. The casting of a shadow on the background in MacLeod's experiment would also be expected, on the basis of contrast, to lighten the disk standing before it.

It is, however, a mistake to conclude that contrast is the *only* factor determining the perception of lightness. Beck, using a different paradigm, found that both cues for illumination and contrast may affect the perception of lightness (1965). The purpose of the present study was to distinguish between the effects due to cues for illumination and those due to contrast, in an experimental setup similar to Gelb's. It was therefore necessary to separate the lightness changes expected on the basis of contrast from the lightness changes expected if cues for illumination affect perceived lightness. In the experiments reported below, a shadow was used to give the impression that a target was strongly illuminated by a spotlight. If changes in lightness are the result of contrast, the presence of a shadow should either lighten the target or produce no change in perceived lightness. If cues for illumination affect perceived lightness, a shadow indicating that a target is in an enhanced illumination should darken the target. In the three experiments reported, the observers judged the lightnesses of two equally illuminated targets, one of which cast a shadow on the background, indicating that the target was illuminated by a spotlight, and one of which did not.

EXPERIMENT I

Method

-Materials–Figure 1 shows a schematic representation of the experimental arrangement. Two pieces of black matte cardboard 4.75 × 4.5 in. were placed on a table 1.25 in. in front of white backgrounds. The walls beyond the immediate white backgrounds were a flat black. The two black surfaces were illuminated by twin projectors set at distances of 72 in. from the surfaces and placed so that they illuminated the surfaces at small slanting angles from the perpendicular. A slide with a circular aperture .25 in. in diameter was inserted into each of the two projectors to produce directed beams of light. The two illuminated areas of the black surfaces served as the targets. Figure 2 shows photographs of the targets with the background shadowed and with the background unshadowed. In the shadow condition the light fell partly on a corner of the surface, and partly on the background above and beside the surface, producing

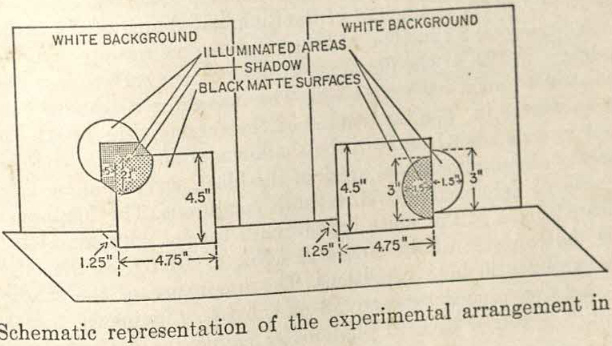


Fig. 1. Schematic representation of the experimental arrangement in Experiment I.

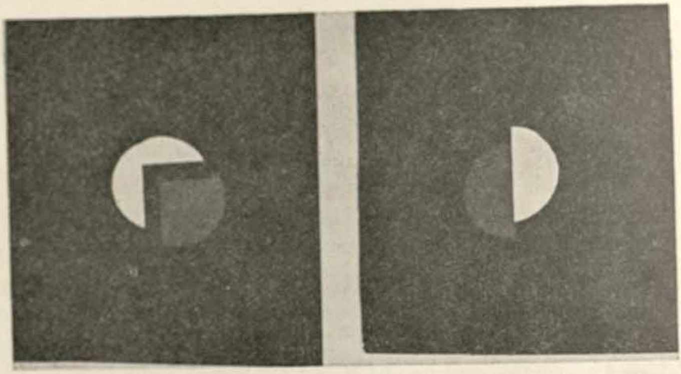


Fig. 2. Photographs of targets in Experiment I. The left picture shows the target casting a shadow on the background, and the right picture shows the target without a shadow. The experimental arrangement is shown in Figure 1.

a deep black shadow with sharp contours. In the no-shadow condition the light fell half on the black surface and half on the white background. The illuminated areas of the black surfaces appeared gray. The illuminated areas of the white backgrounds appeared white. The shadow and the unilluminated areas of the black surfaces and white backgrounds appeared black.

The maximum width and height of the illuminated area of the surface in the shadow condition was 2.1 in. The width of the shadow was approximately .5 in. The maximum widths of the illuminated areas of the surface and background in the no-shadow condition were 1.5 in. The maximum heights of both illuminated areas were 3 in. The luminances of the targets were 14 ftL and of the white backgrounds 180 ftL in both the shadow and no-shadow conditions. The luminances of the unilluminated areas of the black surfaces along the contours of the circles of light were .04 ftL in both conditions. The luminances of the unilluminated areas of the white backgrounds ranged approximately from .42 ftL along the contours of the circles of light, to .04 ftL at the edges of the white backgrounds in both conditions. The luminance of the shadow in the shadow condition ranged between .28 and .5 ftL. Luminance measures were taken with a Spectra-Pritchard photometer employing a 15-min aperture.

-Procedure-Before being brought into the experimental room, an observer was familiarized with a nine-step Munsell scale (white, 9.5, to black, 1.5). The whitest sample was numbered 1 and the blackest 9. The observer was told that in the experiment he would be asked to judge lightness on a scale from 1 to 9. The experiment was conducted in a darkroom. Each observer was brought into the darkroom and stood in front of a screen with a panel set 86 in. from the surfaces. He was told that when the panel was raised he would see two surfaces standing in front of white backgrounds. The targets were always referred to as the left (shadow condition) and the right (no-shadow condition) surface. The observer was instructed to make quick and unanalytical judgments of the lightnesses of the two targets using the lightness scale from 1 to 9 to rate perceived lightness. There were 39 observers in the experiment, and each made a single rating of the lightness of each of the targets in the shadow and no-shadow conditions.

For 24 of the observers a control condition followed the experimental condition. The viewing panel was closed and larger pieces of the same black cardboard were substituted for the surfaces. The light from the projectors now fell completely on the pieces of cardboard. The surfaces under these conditions appeared to be self-luminous and could not be matched to a scale for gray. In the control condition, the observers were asked to judge if the brightnesses of the left and right circles of light were the same or if one was brighter than the other.

Results

Table 1 shows that 29 observers rated the target (illuminated area of the black surface) in the shadow condition as darker and that 10 observers rated the targets in the shadow and no-shadow conditions as of the same lightness. The mean of the observers' lightness matches of the target in the shadow condition was a Munsell value of 7.0 and of the target in the no-shadow condition a Munsell value of 8.1 $[t(38) = 9.09, p < .01]$. The fact that the target in the no-shadow condition was seen as a light gray though

Table 1. Observers' lightness judgments

	Experimental condition			Control condition		
	Shadow condition	No-shadow condition	Same	Left target	Right target	Same
Experiment I						
Target darker	29	0	10	4	3	17
Mean Munsell match	7.0	8.1				
Experiment II						
Target darker	12	1	5	2	0	16
Mean Munsell match	4.6	5.4				
Experiment III						
Target darker	9	3	6	0	4	14
Mean Munsell match	4.3	5.3		7.5	7.3	

Note: The t tests of the difference in mean Munsell matches of the targets in the shadow and no-shadow conditions were significant in Experiments I and II at $p < .01$ and in Experiment III at $p < .05$.

the luminance ratio of the target to the background was approximately 1:13 is consistent with the finding that the darkening produced by an inducing field of higher luminance is decreased if the target and inducing field are perceived at differing depths (Gogel and Mershon, 1969).

In the control condition, 17 observers saw the left and right targets as the same brightness, 4 saw the left target as darker, and 3 saw the right target as darker. The control condition shows that the judged brightnesses of the two targets did not differ when the backgrounds were eliminated.

The fact that a shadow gave rise to an impression of a darker target is inconsistent with a contrast explanation. By that explanation, the shadow, because it was of lower luminance and lay between the black cardboard and the white background, should have decreased contrast and caused the target casting a shadow to be seen as lighter than, or as the same in lightness as, the target without a shadow. The results are consistent, however, with the hypothesis that illumination cues can modify lightness. They support the hypothesis that the shadow provides direct visual information that the target is illuminated by a spotlight. The next two experiments were undertaken to further investigate the effect of a shadow cue on the perception of lightness with somewhat different targets and experimental conditions.

EXPERIMENTS II AND III

Method

-Materials-The aim of Experiments II and III, as of Experiment I, was to dissociate the effects of contrast from those of a shadow cue for illumination. The surfaces in Experiments II and III consisted of 4.75 × 4.5 in. pieces of grosgrain black fabric mounted on stiff cardboard and placed 1.25 in. in front of white backgrounds. The two surfaces were illuminated by two projectors aligned perpendicularly to the surfaces. The luminance variations due to the texture of the fabric introduced internal contrasts that gave the targets (the illuminated areas of the surfaces) a gray appearance and enabled them to be matched to a scale for gray, even in the absence of the white backgrounds. The observers stood on a block so that the shadow behind the target on the left could be clearly seen. Figure 3 shows the appearance of the targets in the shadow and no-shadow conditions.

In the shadow condition the maximum width of the illuminated area of the target was 2.6 in. and the maximum height was 2.3 in. The width of the shadow was approximately .4 in. In the no-shadow condition the maximum width of the illuminated area of the surface was 1.1 in. and the maximum

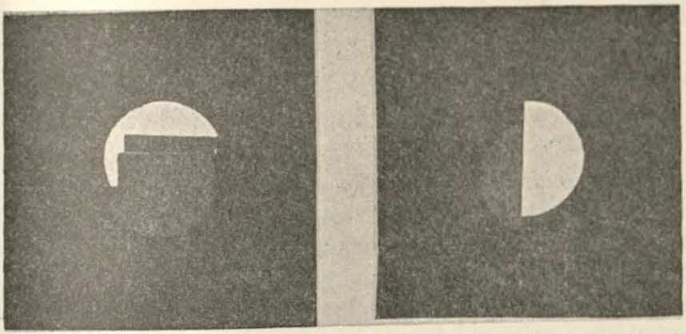


Fig. 3. Photographs of targets in Experiments II and III. The left picture shows the target casting a shadow on the background, and the right picture shows the target without a shadow.

width of the illuminated area of the background was 1.7 in. The maximum heights of both illuminated areas were 3.1 in. The luminances of the targets were .44 ftL and of the white backgrounds 17 ftL in both the shadow and no-shadow conditions. The luminance of the unilluminated area of the target along the contour of the circle of light was .001 ftL in both conditions. The luminances of the unilluminated areas of the white backgrounds ranged approximately from .15 ftL along the contours of the circles of light to .003 ftL at the edges of the white backgrounds in both conditions. The luminance of the shadow was .40 ftL in both conditions.

-Procedure-The two surfaces were alternated so that the surface on the right (no-shadow condition) for one observer became the surface on the left (shadow condition) for the next observer. A control condition in which the circles of light fell wholly on the surfaces was again run following the experimental condition. The remaining procedure in Experiment II was the same as in Experiment I.

In Experiment III, the procedure and apparatus were the same as in Experiment II except that a Munsell chart was placed 78 in. to the observer's left on a stand and illuminated by a 60-w bulb. The bulb was shielded, and a check of the luminances of the targets and backgrounds showed that the addition of the bulb did not alter their luminances. The observers in Experiment III were asked to match the lightnesses of the left and right targets in both the experimental and control conditions to a value on the Munsell scale. Two different groups of 18 observers served in Experiments II and III.

Results

The results of Experiments II and III, shown in Table 1, were like those of Experiment I. The target in the shadow condition appeared darker than the target in the no-shadow condition. In Experiment II 12 observers saw the target in the shadow condition as darker, 1 saw the target in the no-shadow condition as darker, and 5 saw the targets in the shadow and no-shadow conditions as of the same lightness. The mean Munsell match for the shadowed condition was 4.6, and for the no-shadow condition it was 5.4 [$t(17) = 4.23$, $p < .01$].

In Experiment III 9 observers saw the target in the shadow condition as darker, 3 saw the target in the no-shadow condition as darker, and 6 saw the targets in the shadow and no-shadow conditons as the same in lightness. The mean lightness matches of the target in the shadow condition was a Munsell value of 4.3 and of the target in the no-shadow condition a Munsell value of 5.3 [$t(17) = 2.50, p < .05$].

The results of the control conditions show that when the backgrounds were eliminated, the lightnesses of the two targets did not differ significantly. In Experiment II 16 observers saw the left and right as the same lightness, 2 saw the left target as darker, and none saw the right target as darker. In Experiment III 14 observers saw the left and right targets as the same lightness, and 4 saw the right target as darker. The mean Munsell match in Experiment III of the left target was 7.5 and of the right target 7.3 [$t(17) = 1.73, p > .05$].

DISCUSSION

The results of these experiments show that under some conditions, cues for the illumination may affect the perception of lightness. A significant number of observers perceived the target as darker when a shadow was cast on the background than when there was no shadow. The results suggest that the darkening of a target in experimental setups similar to Gelb's may be a function of the stimulus information that a target is illuminated by a spotlight, as well as a function of the inhibitory interactions underlying contrast. It would appear to depend on the salience of the cues indicating that the target is in a special illumination.

Needless to say, the difference between an observer's judgments of the targets in the shadow and no-shadow conditions is not a simple response to information about the conditions of illumination. The observers realized that the targets in the shadow and no-shadow conditions were both illuminated by projectors. Moreover, the brightness of the white background on the no-shadow side would enable the observers to make judgments of the illumination (Beck, 1961). The equal luminances of the white backgrounds would in fact indicate that the two targets were equally illuminated.

Phenomenally, what the shadow produced was the impression of a surface illuminated by a spotlight. Without the shadow the target was seen as a gray surface and the background as a white of a given brightness or insistence. There was no sense of a light falling on the surface, as there was for the target casting a shadow on the background. The perception of the target in the shadow condition as a darker surface appears to have depended on the highly specific information, given through the visual system, that produced the impression of a target illuminated by a spotlight. This is consistent with the earlier finding that lightness perception may depend on specific, visually given stimulus information (Beck, 1969).

Though beyond the scope of the data, a brief discussion of how the shadow cue may influence the processes accounting for the perception of lightness is useful. Beck has proposed (1965) that the perception of lightness involves two separable components: (*a*) the signals resulting from the operation of sensory processes such as adaptation, contrast, and contour formation, which processes produce a central neural pattern in accordance with the peripheral luminance distribution, and (*b*) the integrative schema into which these signals are assimilated. The sensory signals do not specify a particular percept but allow alternative percepts. As a result of past experience, there are stored in memory multiple anchors based on the appearance of black under different illuminations. Whether a stimulus is seen as a black under strong illumination or as a gray depends upon the matching of sensory signals to internal anchors. One may suppose that as a result of neural interactions due to sensory processes, the sensory signals for the targets in the shadow and no-shadow conditions are similar. A difference in perceived lightness results because the shadow cue induces a significant number of observers to assimilate the sensory signals into a schema in which the target is perceived as a darker surface illuminated by a spotlight.

It is important to point out that the apprehension of the illumination is not a necessary condition for lightness constancy. Sensory processes such as contrast and adaptation may operate to keep colors constant with changes of the illumination (Stevens and Stevens, 1960; Hurvich and Jameson, 1966). Moreover, information about the illumination does not in all cases affect the perception of lightness (Beck, 1961). The results of the present and earlier experiments (those by MacLeod, 1932; Kardos, 1934; Gelb, 1929; Beck, 1965) suggest that it is stimulus information which effectively creates the impression of a 'special' illumination—as in the apprehension of a decreased illumination due to a shadow or an enhanced illumination due to a spotlight—that affects lightness perception. It is the perception of a special illumination that appears to be decisive. Thus, a Mach card with one side directly illuminated and one side indirectly illuminated is seen to be of the same lightness on both sides because of stimulus information that causes an observer to apprehend the side of lower luminance as a difference in the illumination of the surface (Beck, 1965).

It is also important to note that basic differences may occur among observers in their attitudes of observation. Henneman reported that observers could make either of two color matches, "taking the standard as 'gray' and ignoring the illumination of the field, or seeing it as 'white' in a region of shadow" (1935, p. 65). The results show that the shadow cue in the present study was not equally effective with all observers in causing the target in the shadow condition to appear darker. Henneman also reported that observers' lightness judgments could be changed by instructions. Informal observations in the present study indicated that some observers could adopt an attitude in which the target in the no-shadow condition could be seen as darker under a spotlight. However, this was generally difficult, and for most observers the shadow cue was necessary for them to see the surface as spotlighted. In any event, the results of the present experiments corroborate earlier findings that lightness perception may be a function of stimulus information concerning the conditions of illumination as well as a function of the inhibitory interactions underlying contrast.

Notes

This research was supported by Grant GB-8027 from the National Science Foundation. Received for publication July 13, 1970.

References

Beck, J. 1961. Judgments of surface illumination and lightness. *Journal of Experimental Psychology* 61:368–373.

Beck, J. 1965. Apparent spatial position and the perception of lightness. *Journal of Experimental Psychology* 69:170–179.

Beck, J. 1969. Lightness and orientation. *American Journal of Psychology* 82:359–366.

Gelb, A. 1929. Die "Farbenknostanz" der Sehdinge. In *Handbuch der normalen und patholgischen Physiologie*, ed. A. Bethe, pp. 594–678. Berlin: Julius Springer, 1925.

Gogel, W. E., and Mershon, A. H. 1969. Depth adjacency and simultaneous contrast. *Perception and Psychophysics* 5:13–17.

Henneman, R. H. 1935. A photometric study of the perception of object color. *Archives of Psychology* no. 179:65.

Hurvich, L. M., and Jameson, D. 1966. *The perception of brightness and darkness*. Boston: Allyn and Bacon.

Kardos, L. 1934. Ding und Schatten: Eine experimentelle Untersuchung über die Grundlagen des Farbensehens. *Zeitschrift für Psychologie* 23.

MacLeod, R. B. 1932. An experimental investigation of brightness constancy. *Archives of Psychology* no. 135.

MacLeod, R. B. 1947. The effects of artificial penumbrae on the brightness of included areas. In *Miscellanea psychologica, Albert Michotte*, pp. 138–154. Paris: Libraire Philosophique.

Newson, L. J. 1958. Some principles governing changes in the apparent lightness of test surfaces isolated from their normal backgrounds. *Quarterly Journal of Experimental Psychology* 10:82–95.

O'Brien, V. 1958. Contour perception, illusion and reality. *Journal of the Optical Society of America* 48:112–119.

Stevens, S. S., and Stevens, J. C. 1960. The dynamics of visual brightness. Report no. PPR-246, Harvard University Psychophysics Project.

Stewart, E. C. 1959. The Gelb effect. *Journal of Experimental Psychology* 57:235–242.

Woodworth, R. S., and Schlosberg, H. 1954. *Experimental psychology*. New York: Holt, 1958.

STIMULUS AND RESPONSE AVAILABILITY IN ASSOCIATIVE LEARNING

Ernest N. Damianopoulos
University of Iowa

A three-stage transfer design (PA1, FRL, PA2) was used to further assess the quantitative contribution of stimulus and response learning to the learning of paired associates. The stimulus and response items were drawn from different populations, and two rates of item exposure (2:2 and 4:4 sec) were employed in the PA2 task. No specific facilitative effects due to stimulus and response learning were observed, contrary to the general conceptualization.

The rather well established generalization that response learning and, to a lesser extent, stimulus learning are component processes in the learning of paired associates was investigated by Damianopoulos (1969) in an effect to assess the *quantitative* contribution of stimulus and response learning (as compared to that of stimulus and response association). Contrary to the earlier confirmatory results of Horowitz and Larsen (1963), Jung (1965), Orlowski and Walsh (1967), and Underwood, Runquist, and Schulz (1959), the results obtained were negative, as no specific facilitative effects were observed in the paired-associates task that followed stimulus and response availability training on a modified free-recall learning task.[1]

Three interpretations were given to these unexpected results. One interpretation was that perhaps the association of stimulus and response plays a much larger role than does the learning of them. The almost perfect terminal recall of the stimulus and response items of the availability training reinforced this interpretation; the items were available but played no role, or at best, a minor one. The second interpretation was that perhaps differential negative and positive effects due to intraserial associations were formed in the availability-training stage and then obscured the effects of stimulus and response learning (Jung, 1967). However, an intraserial-association analysis of the data from the availability training revealed no evidence to support this interpretation. The third interpretation was that perhaps stimulus-response confusion masked any possible facilitative effects, since the CVC sample (the source for both stimuli and responses) had been drawn from a single population; or that perhaps the short 2:2 presentation rate of the paired-associates task masked the facilitative effects, since the items could not be recalled in time.

These three interpretations were tested further in the present experiment. It was hypothesized that if stimulus-response confusion or rate of presentation had contributed to the negative findings of the earlier study, facilitative effects of stimulus and response learning ought to occur when the stimulus and response items were drawn from different populations and a 4:4 rate was added in the paired-associates phase.

METHOD

–Apparatus, materials, and subjects–The apparatus, materials, and subjects were essentially the same as those described earlier (Damianopoulos, 1969). A Lafayette memory drum, Model 303B, was used to display four lists of items. The items in the lists were drawn from two different populations. One-half of the items were CVCs of medium association value (47%–53%) and of ranked difficulty (4–6 on a scale of difficulty from 1 to 9), drawn from a population compiled by Battig.[2] The other items were four-letter English words of medium frequency, from a count compiled by Lorge (1949).

The CVCs and words were arranged in six lists, four of which were for the availability training (free-recall learning; FRL) and two of which were for the paired-associates tasks (PA1 and PA2). The availability-training lists consisted of two five-item CVC lists and two five-item word lists, designated A, B, C, and D: A and C were word lists, B and D were CVC lists. The paired-associates lists were made up of word-CVC pairs, the word as stimulus, the CVC as response item. One of the lists, A-B, was derived from paired ('relevant') A and B items; the other list, E-F, consisted of items unrelated ('irrelevant') to the availability-training list.

The subjects were volunteers from the introductory courses at Indiana University Southeast and were assigned in order of appearance to 10 experimental and control groups, 10 subjects to each group. Four subjects were dropped and replaced: two because they failed to learn a three-item practice list, the others because they took more than 30 trials to reach criterion on one of the lists of paired associates.

–Design and procedure–The three-stage transfer design consisted of an irrelevant paired-associates task (E-F), a *paced* availability-training task that included relevant and irrelevant items, and finally a relevant paired-associates task (A-B). The decision to place the irrelevant paired-associates task first was guided by two considerations: (*a*) a need to assess directly the equality of the groups prior to administration of the main treatment and (*b*) a need to reduce the variance attributable to learning-to-learn effects prior to the main treatment (so as to increase design precision). Altogether, there were 10 treatment groups arranged in a 2 × 5 factorial design. One dimension consisted of five treatment conditions in availability training (A,B; A,D; C,B; C,D; and O, and the other dimension consisted of two rates of item presentation in the A-B task (2:2 and 4:4 sec).

After mastering a three-item practice list, the subject learned the E-F list. Upon mastering this task, he learned two availability-training lists in sequence. Depending on his group assignment, he learned either the A,B lists, the A,D lists, the C,B lists, the C,D lists, or, in Group O, nothing. In the practice trials of availability training, the items were presented in a random sequence at a 2-sec rate for all groups. In the test sessions, the subject, upon observing a signal to start (two red arrows), recalled, in any order, the items from the list, again at a 2-sec rate. The memory drum was used here as a *pacing* device. Every time the memory drum advanced the empty paper scroll, the subject was required (by the instructions) to recall an item from the list. In this way, the verbal responses were paced to occur only during the 2-sec interval when the paper scroll was on display. The subjects in Group O were instructed to sit for 5 min on a nearby chair, until the experiment could be resumed. Half of the subjects learned the CVC list first, then the word list; while the other half learned the lists in the reverse order.

The A-B list of paired associates was administered immediately upon completion of availability training. There was a 2:2-sec presentation rate for one group, and a 4:4 rate for the second group. All these tasks were learned to a criterion of one perfect trial. There was, throughout, a 1-min intertask interval. At the end of the experiment, the subject was asked to recall the stimulus and response items of the availability-training task and also to describe how the two paired-associates tasks were learned (mnemonic devices). Finally, the subject was thanked for participating in the experiment and cautioned not to discuss it with any but those who had already participated.

RESULTS

The mean number of trials required to learn E-F, the first paired-associates task, to a criterion of one errorless trial ranged from 5.9 to 12.8 trials. The treatment groups A,B, A,D, C,B, C,D, and O of the 2:2 rate (at this point these are all dummy groups) required 10.3, 5.9, 11.9, 9.0, and 7.9 trials respectively, while the corresponding groups of the 4:4 rate required 12.8, 10.3, 6.5, 12.1, and 11.2 trials respectively. A two-factor orthogonal analysis of variance revealed that none of the mean differences reached significance, either for treatment [$F(4, 90) = .754, p > .05$], or for rate [F(1,

90) = 1.359, $p > .05$], or for the interaction of treatment and rate [$F(4, 90) = 1.646$, $p > .05$]. A correlational analysis of the relationship of performance on E-F to that on A-B was carried out to investigate the assumptions that performance on E-F was an adequate indicator of treatment-group equality for the criterion task, A-B, and that E-F learning could indeed reduce variance, due to learning-to-learn effects, on task A-B. Separate analyses for the 2:2 and 4:4 groups revealed a moderate positive relationship, +.39 and +.64 respectively (Lindquist, 1953, p. 321).

The number of trials required to learn the two availability-training lists to a criterion of one errorless trial appears in Table 1, which shows that the mean number of trials for the five-item CVC lists ranged from 3.3 to 6.3 trials. A two-factor orthogonal analysis of variance revealed no significant difference among the mean differences for treatment [$F(3.72) = 2.153, p > .05$], for rate [$F(1, 72) = .004$ $p > .05$; the rate groups here again being dummy groups], or for the interaction of treatment and rate [$F(3, 72) = .600, p > .05$]. As can be seen from Table 1, the five-item word lists were learned within two trials. A similar analysis of variance revealed that the treatment means (ignoring presentation rate) differed significantly [$F(3, 72) = 4.26, p < .05$]. However, no practical significance could be attached to this finding since the overall treatment means *ranged* only from 1.2 to 2.4 practice trials. All other comparisons were nonsignificant.

Table 1. Mean number of practice trials to one perfect trial in availability training

	Group							
	A,B		A,D		C,B		C,D	
Presentation rate	CVCs	Words	CVCs	Words	CVCs	Words	CVCs	Words
2:2-sec rate								
$\bar{X}$	6.3	1.2	3.5	2.2	4.6	1.3	3.3	1.1
SD	3.7	.4	1.2	1.9	2.9	.4	1.4	1.1
4:4-sec rate								
$\bar{X}$	5.2	1.2	4.2	1.3	4.5	1.1	4.3	2.4
SD	1.6	.4	2.0	.4	3.3	.8	1.4	.8

Note: Here as elsewhere 'rate' refers to the presentation rate on list A-B.

The specific effects of stimulus and response availability on the learning of the paired associates were determined by comparing

Groups A,B, A,D, and C,B against Group C,D on the criterion measure, number of trials to learn A-B. It was expected that, with differentiated stimulus and response items and with the added 4:4-sec rate, Groups A,B, A,D, and C,B (as compared to Group C,D) would show facilitation at both presentation rates. However, as Table 2 shows, the mean numbers of trials to criterion were virtually identical for all groups at either rate. A two-factor orthogonal analysis of variance revealed no overall significant differences among the treatment means (ignoring presentation rate) [$F(4, 90) = .262$, $p > .05$]. Presentation rate (ignoring treatment), however, did show an overall difference [$F(1, 90) = 27.711$, $p < .01$]. The 4:4 groups required fewer trials than did the 2:2 groups, which is in keeping with prior studies showing that presentation rate is an effective variable. The expected interaction of treatment and presentation rate did not materialize [$F(4, 90) = .192$, $p > .05$]. A covariance analysis using, for each subject, the number of trials to criterion on E-F as the control measure similarly revealed no differences among the treatment means [$F(4, 90) = .284$, $p > .05$]; it did reveal a clear-cut effect of presentation rate [$F(1, 90) = 39.969$, $p < .01$]; again, the interaction was not significant [$F(4, 90) = .414$, $p > .05$]. Thus, the results neither show facilitative effects due to stimulus and response availability nor do they demonstrate nonspecific effects at either the short, 2:2, or the longer, 4:4, presentation rate.

Table 2. Mean number of practice trials to one perfect trial on List A-B

	Group				
Presentation rate	A,B	A,D	C,B	C,D	O
2:2-sec rate					
$\bar{X}$	12.1	10.0	11.5	11.3	11.6
SD	3.5	3.3	6.1	7.2	5.1
4:4-sec rate					
$\bar{X}$	6.3	6.2	6.1	5.2	7.1
SD	4.1	2.8	5.9	1.7	3.6

Since in a Task 1–Task 2 transfer paradigm, transfer effects are mainly confined to the first few learning trials of the second task, the learning rates on the first five trials of the 2:2 groups and the first three trials of the 4:4 groups were examined, to see if any effects of stimulus and response availability might show up there.

Table 3 shows the mean numbers of items learned. A Lindquist (1953) Type I analysis of variance was performed separately for the two presentation rates. It revealed that treatment mean differences at the 2:2 rate did not reach significance [$F(4, 45) = .763$, $p > .05$]; practice-trial mean differences were significant [$F(4, 180) = 21.847$, $p < .01$]; the interaction of treatment and trials was not significant [$F(16, 180) = .767$, $p > .05$]. The same analysis at the

Table 3. Mean number of correct items on the early trials on List A-B

	Trials				
Presentation rate	1	2	3	4	5
2:2-sec rate					
Group A,B					
$\bar{X}$	1.0	1.1	1.5	2.4	2.5
SD	1.5	.7	1.5	1.0	.7
Group A,D					
$\bar{X}$	1.3	1.5	1.9	2.4	3.0
SD	.6	1.0	1.0	.9	.8
Group C,B					
$\bar{X}$	1.8	1.7	2.1	2.7	3.3
SD	1.2	1.6	1.5	1.1	1.0
Group C,D					
$\bar{X}$	1.7	1.9	2.8	2.7	3.1
SD	.4	.5	.4	.4	4.
Group O					
$\bar{X}$	1.8	1.9	2.1	2.7	2.3
SD	.5	.5	.5	.4	.5
4:4-sec rate					
Group A,B					
$\bar{X}$	2.2	3.0	3.7		
SD	1.6	1.3	1.3		
Group A,D					
$\bar{X}$	1.3	2.4	2.9		
SD	.9	1.5	1.3		
Group C,B					
$\bar{X}$	2.4	3.0	3.7		
SD	1.1	1.3	1.3		
Group C,D					
$\bar{X}$	.9	2.3	3.3		
SD	.9	.8	1.6		
Group O					
$\bar{X}$	2.0	2.7	3.4		
SD	.6	1.0	.9		

4:4 rate similarly revealed no significant treatment mean differences $[F(4,45) = 1.972, p > .05]$, a significant effect of trials $[F(2,90) = 33.367, p < .01]$, and no significant interaction of trials and treatment $[F(8,90) = .758, p > .05]$. A covariance analysis using, for each subject, the number of trials to criterion on E-F as the control measure revealed no differences among the treatment means on the total number of items learned after five trials $[F(4,45) = .736, p > .05]$. A similar analysis for the 4:4 groups after three practice trials also revealed no treatment mean differences $[F(4,45) = 1.912, p > .05]$. Thus, once more, there were no observable effects of stimulus and response availability on the learning of paired associates.

An analysis of the data on the CVCs of availability training was carried out to determine if the results described above could be attributed to that training's differential intralist-association effects on the A-B task (Jung, 1967). The number of items occurring in an adjacent relation, forward or backward, in the next-to-the-last and in the last trials were compiled. The results are presented in Table 4. The individual means shown were then compared by a *t*-test analysis against an expected (chance) mean of one adjacent (forward or backward) relation. None of the mean comparisons reached significance according to conventional *t*-test analysis. The expected mean was determined empirically. Twenty sequences of numbers 1–5 were recorded from a table of random numbers, repeated numbers being excluded. Taking two sequences at a time, a count was then made of repeated two-number sequences (i.e., numbers adjacent in either forward or backward direction). Another restriction was imposed in that on $n - 1$ of each set, the last number was always discarded. The observed rate of adjacent-

Table 4. Mean number of items occurring in an adjacent (forward or backward) relation in the next-to-last and last trials of availability training

	Group			
Presentation rate	A,B	A,D	C,B	C,D
2:2-sec rate				
$\bar{X}$	1.2	1.0	.8	.8
SD	.8	1.3	.8	.6
4:4-sec rate				
$\bar{X}$	.5	1.2	.5	.5
SD	.5	1.0	.7	.7

number frequency was 2, 1, 0, 1, 2, 0, 1, 1, 0, and 1 for the 10 sets of five-number sequences.

An analysis of overt errors on list A-B was carried out as an alternate method of assessing intralist associations from availability training. For each subject a count was made of the total number of misplaced responses made in learning list A-B to criterion. The results appear in Table 5, which shows that the group means ranged from 2.9 to 9.7. A two-factor orthogonal analysis of variance revealed no significant overall treatment mean difference [$F(4, 90) = 1.222, p > .05$]. The presentation-rate comparison did reveal a marginal difference between means [$F(1, 90) = 2.793, p < .05 > .025$]; however, following Linquist's advice (1953, p. 83) on violations of the homogeneity-of-variance assumption, a stricter alpha criterion ($p \leq .025$) was adopted and the null hypothesis could not be rejected. The interaction of treatment and rate was not significant [$F(4, 90) = .559, p > .05$]. Thus, again, no affirmative evidence was observed for possible differential intralist-association effects on the criterion task, A-B.

Table 5. Mean number of misplaced responses on List A-B

	Group				
Presentation rate	A,B	A,D	C,B	C,D	O
2:2-sec rate					
$\bar{X}$	6.0	3.3	9.7	6.7	6.9
SD	3.5	2.9	8.5	8.3	5.0
4:4-sec rate					
$\bar{X}$	3.9	4.3	5.7	2.9	5.6
SD	3.8	3.2	9.4	2.2	5.4

A terminal recall test of the CVCs and words showed, in contrast to the results of the previous study, that the items from availability training could *not* be readily recalled. The highest rate of recall was on the relevant items (those which appeared as stimulus and response items in the subsequent A-B task) and it reached only 50%.

DISCUSSION

Consistent with the previous findings by Damianopoulos (1969), the present results failed to reveal specific facilitative effects of stimulus and response availability on a subsequently learned paired-associates task. No effects were observed either on the criterion measure—number of trials to learn list A-B—or on the number of items learned in each of the early practice trials on list A-B. These results are contrary to a considerable body of evidence that response learning and, to a lesser extent, stimulus learning are component processes in associative learning (Horowitz and Larsen, 1963; Jung, 1965; Orlowski and Walsh, 1967; Underwood, Runquist, and Schulz, 1959).

The present results also failed to support two of the interpretations hypothesized to account for the earlier negative results (Damianopoulos, 1969). There was no evidence of differential intralist-association effects on the A-B task. Differentiating the stimuli from the responses in availability training and in the A-B task did not lead to a facilitative effect; nor did lengthening the item rate of presentation from 2:2 to 4:4 sec in the A-B task. Thus, the only remaining interpretation seems to be that stimulus and response learning (as compared to the association of them) play no role at all, or at best, a minor one, as component processes in the learning of paired associates. Perhaps one can conceptualize stimulus and response learning as secondary processes, or even as by-products, of the process of stimulus-response association. If the paced-recall trials during availability training prevented rehearsal and other organizational processes from taking place in the present experiment, one could cite this factor as a possible explanation of the earlier positive findings.

The findings of the present experiment are not as unique as they might first appear. For example, Blanchard, using a similar design, found that a group receiving 40 familiarization trials on irrelevant items did not differ significantly from a group that got 40 trials on relevant items (1967). Since the first group was the proper control for evaluating *specific* transfer effects due to response familiarization, it can be concluded that Blanchard, too, did not find evidence for the response-learning hypothesis. Blanchard was "puzzled" (p. 670), but the message remained uncommunicated. Nelson also did not find *specific* transfer effects (1967). These two investigators apparently failed to consider it an operational requirement that only specific transfer effects be considered as indicative of stimulus- and response-learning effects on a subsequently learned paired-associates task.

The conclusion reached above—based, as it is, on negative

BUREAU OF EDUCATION LIBRARY

findings—is open to at least two criticisms. The first stems from the relative ease with which the availability-training task was learned. Since the five-item response list was learned in about five trials and the five-item stimulus list was learned in only one or two trials, the task might have been too easy to have had an effect on the subsequently learned paired-associates task. The second criticism stems from the moderate correlation observed between E-F and A-B list learning. The inference of treatment-group quality on A-B from E-F performance is open to doubt, as is the inference about learning-to-learn effects. One explanation of the correlation, however, might be based on the progressive effects of the first task on the second task. Inasmuch as the tasks were sampled from the same source, the subjects who required a large number of trials to learn the first task might, on that basis alone, have required a relatively small number of trials on the second task. And learning-to-learn effects from the first task might have been greater for the subjects who spent a longer period on the first task; hence, the moderate correlation.

In view of the two criticisms, and others concomitant with negative results, the effects of stimulus and response learning on the learning of paired associates might be evaluated further by such an alternative method of producing stimulus and response availability as using paired items (n-S, n-R), as tasks prior to the relevant (S-R) paired-associates task.

Bureau of Ednl. & Psyl. Researc
(S. C. E. R. T.)
Date 3.3.81
Acc J652

Notes

Received for publication June 12, 1970.

1. A distinctive feature of the Damianopoulos (1969) study was that the availability training was administered on relevant and irrelevant stimulus and response items by means of a *paced* free-recall learning task where the item-presentation rate was held constant at 2 sec with a 2-sec interval.

2. American Documentation Institute 5847.

References

Blanchard, R. S. 1967. Response familiarization and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior* 6:668–671.

Damianopoulos, E. N. 1969. Effect of stimulus and response availability on associative learning. *Journal of Experimental Psychology* 80:572–574.

Horowitz, L. M., and Larsen, S. R. 1963. Response interference in paired associate learning. *Journal of Experimental Psychology* 65:225–232.

Jung, J. 1965. Two stages of paired associate learning as a function of intralist response similarity (IRS) and response meaningfulness (M). *Journal of Experimental Psychology* 70:371–378.

Jung, J. 1967. Transfer analysis of familiarization effects. *Psychological Review* 74:523–529.

Lindquist, E. F. 1953. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin.

Lorge, I. 1949. *The semantic count of the 570 commonest English words.* New York: Teachers College, Columbia University.

Nelson, D. L. 1967. The effect of prior response learning upon subsequent paired-associate learning under constant or varied orders of retention. *Journal of Verbal Learning and Verbal Behavior* 6:78–83.

Orlowski, W. J., and Walsh, J. F. 1967. The effect of stimulus familiarization procedure on paired-associate verbal learning. *Psychonomic Science* 8:435–436.

Underwood, B. J., Runquist, W. A., and Schulz, R. W. 1959. Response learning in paired-associate lists as a function of intralist similarity. *Journal of Experimental Psychology* 58:70–78.

A COMPARISON OF TWO METHODS OF VALUE MEASUREMENT

Theodore W. Lorei and Jacob Cohen
Veterans Administration Hospital, Washington, D.C., and New York University

Category rating and ratio estimation were used by 621 staff respondents to scale the importance (value) of 16 possible outcomes of releasing psychiatric patients or keeping them in the hospital. Stevens's claim that the relationship between the two methods is intermediate between linear and logarithmic was confirmed for seven of the eight occupational groups—but not for psychologists. The two methods yielded essentially identical answers to the substantive survey questions.

The writers have recently reported the results of a multi-hospital survey of staff opinion about the importance (value) of several possible outcomes of releasing psychiatric patients or retaining them in the hospital (Lorei, 1970; Lorei and Cohen, 1970). The present article focuses in detail on the procedures used to scale these judgments of importance and compares the results of two different scaling procedures.

The survey objectives required that the importance of the outcomes be measured on an interval scale, and according to common opinion, several procedures—pair comparisons, category rating, ratio estimation—should produce such scaling. Recent research, however, has shown that for prothetic continua[1] the procedures just mentioned do not yield equivalent results: the scale values are not linearly related (Ekman and Sjöberg, 1965; Stevens, 1966). The usual finding is that when a category or pair-comparison scale is plotted along the ordinate and a ratio-estimation scale along the abscissa of linear coordinates, the curve is 'concave downward.' Stevens interprets this relation as reflecting unequal intervals on the pair-comparison and category scales (1966).

Assuming (*a*) that neither pair comparison nor category rating yields interval-scale measurement for prothetic attributes and (*b*) that importance (value, utility) is a prothetic variable, as Stevens suggested (1959, p. 53), ratio estimation would seem to be the procedure of choice in the present context. Not all are as convinced as Stevens, however, that the nonlinearity between the results of the scaling procedures is due to bias in the pair comparisons and category ratings (see Ekman and Sjöberg, 1965). Further, a pilot study of hospital staff showed an extremely high (linear) correlation between scale values resulting from category rating and direct estimation, suggesting that it made little difference which one was used. (The pair-comparison technique was not considered because the relatively large number of outcomes would have made the comparison task unduly burdensome.)

The present paper reports a comparison of the results of using both category-rating and ratio-estimation methods with a large sample of hospital staff. More specifically, the study sought to determine (*a*) whether the form of the relation between the results of the two scaling procedures was consistent with the results reported by Stevens (1966) and (*b*) whether the two methods yielded essentially identical answers to the substantive survey questions.

METHOD

–Category rating–Sixteen possible outcomes of releasing or retaining patients were formulated and presented to hospital staff in the form of the Release/Retention Outcome Inventory. The inventory items appear in Table 1. Subjects were instructed to rate the importance of each outcome on a scale ranging from 0 (not at all important) to 8 (extremely important). Although they were not explicitly told to consider the intervals between scale points as equal, it was assumed that the scale-point numbers would imply such equality to them. To establish a frame of reference that would not shift during the rating process, the subjects were asked to look over the total set of outcomes before beginning.

–Ratio estimation–Essentially the same items, with very minor rewording necessary in some cases to fit the rating format, were presented to the subjects in a diagrammatic form to elicit ratio judgments.[2] The diagram was a flow chart showing a branch point labeled "to release a mental patient or not" followed by two blocks labeled "release" and "keep." The possible release and retention outcomes were listed after the appropriate block in a column captioned "possible outcomes in the future." Subjects were instructed to read through the list of possible outcomes, select the most undesirable outcome, and assign it a value of —100. They were further instructed to rate the

Table 1. Items of Release/Retention Outcome Inventory

If it were entirely up to you whether or not to release a patient, *how important would it be to you that . . .*

1. He be recovered enough so that he wouldn't cause his family a lot of worry and annoyance?
2. He be recovered enough so that people wouldn't be able to take advantage of him (such as cheating him out of his money or not paying him enough for his work)?
3. He be recovered enough so that people wouldn't make fun of him or avoid him?
4. He spend his time working or doing something else useful (for example, helping out at home or doing volunteer work)?
5. He be out of the hospital so he could decide things for himself and run his own life?
6. He not seriously hurt (or even kill) himself or anybody else?
7. His family not complain to important people (such as a congressman) because he was released? (This question assumes that, in your best judgment, he should be released.)
8. He not be very lonely or unhappy outside of the hospital?
9. He not drink so much that he would injure his health, get into fights, or make life miserable for others?
10. He not be a serious financial or physical burden on his family?
11. He not be too unhappy with your decision, whichever way it went?
12. He not need to seek support from "welfare" for the necessities of life (food, clothing, etc.)?

In deciding whether to release the patient or not, you would also want to consider some of the possible results of *keeping* him in the hospital. How important would it be to you that . . .

13. He not become "institutionalized," that is, that he would not become withdrawn and uninterested in leaving the hospital?
14. He not use the hospital mainly as a hotel and thus avoid going to work or facing other responsibilities?
15. You not be criticized by "higher-ups" because the patient turnover on your ward is too low to satisfy them?
16. His family not complain to important people (such as a congressman) about your keeping the patient in the hospital? (Again, it is assumed that, in your best judgment, he really needs to be there.)

Note: Immediately following each item, a scale ranging from 0 to 8 was presented with the following scale-point definitions: not at all important, very slightly, a little, mildly, moderately, quite a bit, distinctly, markedly, and extremely important.

undesirability of the remaining outcomes by assigning numbers that reflected their ratios to the outcome rated −100.

-General procedure and subjects-For both sets of ratings the subjects were asked to assume that release/retention decisions were entirely up to them, whether or not their usual responsibilities included this kind of decision making. The two sets of ratings were obtained in an individual session with a research psychologist or social worker. The inventory was always administered first so that whatever order effects might exist would be uniform. Subjects were asked to read the instructions for both instruments, and they were encouraged to ask for further explanation if they felt the need. After both sets of ratings had been completed, the interviewer asked a few questions to determine how well the subject had understood the rating task. Both sets of ratings were obtained from 621 subjects from 6 of 12 participating hospitals. The ratings were obtained from the eight occupational groups listed in Table 2. Random samples of nurses and nursing assistants were selected; all eligible employees in the remaining six groups were included. Only the inventory was administered at the remaining 6 hospitals.

The mean ages of the occupational groups ranged from 40 for nursing assistants to 56 for physicians. The mean number of years of education varied from 11 for nursing assistants to 20 for psychiatrists, other physicians, and psychologists.

RESULTS

Understanding of rating tasks

For the 621 respondents who performed both scaling tasks, the percentages of the subjects rated as 'very probably meaningfully completing' the task were 74.5 for category rating and 67.1 for ratio estimation; and as 'quite probably meaningfully completing,' 18.7 for category rating and 24.0 for ratio estimation. The differences are significant and, when netted out, favor the category rating, as anticipated, but not to a substantial degree. Thirteen subjects were dropped from further analysis because their understanding or cooperation was judged very poor.

Relation between category and ratio scaling

To determine the degree and form of the relation between the two scales, the following analyses were performed. The scale values (arithmetic means) for each of the 15 comparable outcomes resulting from each procedure were computed for each of the eight occupational groups and for all groups combined. Stevens states that the relationship for prothetic continua is not linear but, rather, "roughly" logarithmic (1968, p. 179) or, more specifically, "intermediate between linear and logarithmic" (1968, p. 177). To test this proposition, two series of polynomial multiple-regression analyses (Cohen, 1968) were run on each group's scale values from the two

Table 2. Results of two series of polynomial multiple-regression analyses for relating category-rating and ratio-estimation scale values of outcomes

Independent variables	Cumulative R^2	ΔR^2	Independent variables	Cumulative R^2	ΔR^2
		Psychiatrists ($N = 64$)			
RE	.935[a]	.935**	LRE	.964	.964**
RE^2	.971	.036**	LRE^2	.973	.009
RE^3	.971	.000	LRE^3	.973	.000
		Physicians ($N = 26$)			
RE	.868	.868**	LRE	.908	.908**
RE^2	.917	.049*	LRE^2	.920	.012
RE^3	.918	.001	LRE^3	.920	.000
		Nurses ($N = 179$)			
RE	.928	.928**	LRE	.955	.955**
RE^2	.966	.038**	LRE^2	.965	.010
RE^3	.966	.000	LRE^3	.968	.003
		Nurses assistants ($N = 165$)			
RE	.932	.932**	LRE	.973	.973**
RE^2	.975	.043**	LRE^2	.974	.001
RE^3	.975	.000	LRE^3	.974	.000
		Psychologists ($N = 46$)			
RE	.982	.982**	LRE	.935	.935**
RE^2	.985	.003	LRE^2	.983	.048**
RE^3	.985	.000	LRE^3	.984	.000
		Social workers ($N = 64$)			
RE	.945	.945**	LRE	.926	.926**
RE^2	.967	.022*	LRE^2	.964	.038**
RE^3	.969	.002	LRE^3	.965	.001
		PMR personnel ($N = 41$)			
RE	.912	.912**	LRE	.937	.937**
RE^2	.955	.043**	LRE^2	.951	.014
RE^3	.955	.000	LRE^3	.953	.002
		Administrative personnel ($N = 23$)			
RE	.925	.925**	LRE	.910	.910**
RE^2	.956	.031*	LRE^2	.953	.043**
RE^3	.961	.005	LRE^3	.960	.006
		Total group ($N = 608$)			
RE	.943	.943**	LRE	.959	.959**
RE^2	.976	.033**	LRE^2	.973	.014*
RE^3	.976	.000	LRE^3	.973	.000

Note: PMR personnel are Physical, Medicine, and Rehabilitation personnel.
[a]All cumulative R^2 in this table significant at $p < .01$. $^{**}p < .01$. $^{*}p < .05$.

procedures to determine the form of the relationship between them.

First, to test the hypothesis of nonlinearity, using category rating as the dependent variable, analyses were performed using as independent variables the ratio estimations RE, RE^2, and RE^3 cumulatively (i.e., first RE, then RE and RE^2, then RE, RE^2, and RE^3). The first three columns of Table 2 give for each group the R^2 at each stage, together with the change from the previous stage (Δ), and the significance of the latter.

Second, to test the hypothesis that the nature of the relationship is logarithmic or between logarithmic and linear, analyses were performed again with the category-rating scale value as the dependent variable, but now with the logarithm of the ratio-estimation scale values as independent variable used as a polynomial set, as above (e.g., LRE, LRE^2, LRE^3). The last three columns of Table 2 give the R^2 and Δ at each stage for each group, as before.

From Table 2, it is apparent that the linear relationship between ratio-estimation and category-rating scale values found in the first series of analyses was strong for the ($N = 15$) rated outcomes. The squared correlations ranged from .868 (physicians) to .982 (psychologists), with a median of .934. (In the more familiar terms of r, the values were .932 to .991, with a median of .966.) Granted the strength of this linear relationship, the addition of RE^2, which provides for simple ("one-bend"; see Cohen, 1968) curvilinearity, brought a significant improvement in all the groups except for the psychologists (to whom we will return below). With the same exception, the regression coefficient for RE^2 was always negative, indicating that the regression of category rating on ratio estimation was concave downward. Moreover, this relationship had no hint of second bend: the ΔR^2 for RE^3 was uniformly trivial. Thus, in seven of eight groups the relationship was curvilinear and of the general shape Stevens posits.

If we compare the squared correlations with category-rating values for ratio-estimation values (from the first series of analyses) with those for the logarithm of the ratio-estimation values (from the second series of analyses), we find the eight groups equally divided as to whether the linear or logarithmic relationship was the larger. This is consistent with Stevens's claim of a relationship whose form is intermediate between the two. When the form of the regression of category-rating values on the logarithm of ratio-estimation values is further studied by adding LRE^2, in three of the eight groups the increase in R^2 was significant (with only 12 error *df*), and in all eight groups the regression coefficient for LRE^2 was positive. Therefore, the regression of category-rating values on the logarithm of ratio-estimation values was always concave upward, indicating that the departure from a logarithmic function was always in the linear direction, although not necessarily significantly so. Again, there was no hint of a double bend from ΔRE^3.

Taken in all, in seven of the eight groups, Stevens's claims were exactly supported: the regression of category-rating on ratio-estimation values was nonlinear, in fact lying precisely between a linear and logarithmic form. The psychologists' scalings constituted a perverse exception. For them, the category rating–ratio estimation relationship was emphatically linear: the squared linear correlation (from the first series of analyses) was .982 (hence, $r = .991$) and the addition of RE^2 yielded virtually no increase; indeed, there was hardly room for any increase. Consistent with this, the squared correlation of category-rating values with the logarithm of ratio-estimation values (from the second series of analyses) was only .935; only when LRE^2 was added, and bent the logarithmic regression back to linearity, did R^2 attain the same magnitude (.984) as the squared linear correlation (.982). Thus, the psychologists uniquely among the eight occupation groups showed an almost perfect linear relationship between their category-rating and ratio-estimation scalings of the 15 outcomes. We can only speculate about this peculiar phenomenon. Perhaps the training of psychologists sensitizes them to avoid permitting variability to increase with the intensity of the prothetic continuum, the reason given for the nonlinear relationship (Ekman's law; see Stevens, 1968). Perhaps this particular continuum—importance of outcomes—was not intensive (prothetic) but, rather, qualitative (metathetic) for the clinical psychologists making these ratings, in which case the relationship of the two methods should be linear (Stevens, 1968, p. 177). Or, finally, this may simply have been a chance departure toward linearity from the usual relationship. Only further scaling research with psychologists as subjects can solve this puzzle.

In brief, Stevens's claim that the relationship between category and ratio (or magnitude) scaling of attitudes is intermediate between linear and logarithmic was strikingly confirmed for seven of the eight occupational groups rating the importance of outcomes of the release/retention decision in psychiatric hospitalization.

Consequences of the two scaling procedures for the substantive survey questions

The question of how the results of the two scaling procedures relate to each other is an important one for theory. The second issue of this paper concerns the differences that the use of one or the other procedure would make for a specific *substantive* application. To investigate this second point, the analyses bearing on the two central questions of the outcome-value survey were run using both the category-rating and the ratio-estimation data.

The survey was concerned with determining how hospital personnel *as a whole* scale the relative importance of release/retention outcomes. As indicated in Table 2, the squared linear correlation of the 15 pairs of outcome means for all subjects combined was .943, hence $r = .971$. It was accordingly concluded that the scaling procedure used made little difference here.

The survey also sought to determine whether occupational groups differed in their scaling of outcome values, and if so, by how much. An analysis of variance for repeated measurements on several independent groups was performed on both the category-rating and the ratio-estimation data. The results are presented in Table 3. What is chiefly of note in this table is that all three nonerror sources of variance yielded larger F ratios and hence ϵ^2 (Cohen, 1965)[3] for the category-rating scaling than for the ratio-estimation scaling, although in two of these the difference was relatively small. Since all three were significant by both scaling methods, the same conclusions would be drawn.

The substantive issue of most interest in Table 3 is the mean square of the interaction of outcomes and groups. If the groups' value profiles[4] were perfectly parallel, this term would be zero. The significant F ratios indicate that occupational groups did scale outcome value differently, but the ϵ^2 values indicate that these profile differences were small (see Cohen, 1965). More precisely, the ϵ^2 can be interpreted as the proportion of variance in individual profiles that is explained by group membership, that is, by differences in *group* profiles.

The between-groups mean squares in Table 3 indicate that the occupational groups differed in the mean levels of their outcome profiles or, alternatively, that the groups differed in the extent to which they valued the outcomes regardless of the specific content. The associated ϵ^2 suggests that the category ratings were much

Table 3. Analysis of variance of outcome values

Source	*df*	*MS*		*F*		ϵ^2	
		Ratio estimation	Category rating	Ratio estimation	Category rating	Ratio estimation	Category rating
Between subjects	607	28.32	15.48				
Between groups	7	230.18	228.55	8.87**	17.59**	.08	.16
Between subjects within groups	600	25.96	12.99				
Within subjects	8512	10.41	6.77				
Between outcomes	14	3335.93	2197.58	675.29**	720.52**	.53	.54
Outcomes × subjects	8498	4.94	3.16				
Outcomes × groups	98	13.38	12.42	2.77**	4.07**	.02	.03
Outcomes × subjects within groups	8400	4.83	3.05				

$^{**}p < .01$.

more sensitive to level differences than the ratio estimates or, perhaps, that category ratings were more subject to the rater's frame of reference or to context effects than are ratio estimates.

The between-outcomes mean square reflects the sharpness of the separation of values (i.e., scaling). The associated ϵ^2 indicates that the results of the two scaling procedures in this regard were virtually identical, the slight difference favoring the category-rating procedure.

DISCUSSION

The relation between the outcome-value scales resulting from category ratings and ratio estimates is fully consistent with the results of previous research, as well as with Stevens's theorizing. While Stevens's hypothesis is that the nonlinearity of this relation results from bias in the category rating (1966), this is by no means a universally accepted view (see Ekman and Sjöberg, 1965). At any rate, for practical purposes it seems to make little difference which procedure is used. If, however, it is true that ratio estimates are less influenced by context than are category ratings, as is claimed by Galanter (1962, pp. 152–153), then ratio estimates might be more desirable, particularly if new release/retention outcomes were to be added to the present set.

The growing interest in using decision theory to suggest choices in industry, business, mental health, and many other areas has made the measurement of value, a key decision-theory concept, of considerable practical importance. Further, while even a simple ranking of outcomes in terms of value can have considerable usefulness, interval and ratio measurements permit a much fuller exploitation of value data.

It seems reasonable to believe that in the present study, which took place within a field setting, interval and even ratio measurement was achieved. Whether one believes that such measurement has been achieved, or is even possible, would seem to be a function of one's measurement philosophy or, in Coombs's phrasing (1964, p. 284), of one's preference for strong or weak measurement models. Coombs himself, despite his acknowledged general preference for weak models, feels that society's need for decision and action provides a persuasive argument for making strong assumptions and using strong models in value measurement. The basic implication of the present study and of the survey it reports is that relatively refined value measurement of outcomes relevant to everyday organizational decision making can be achieved and should be attempted more often.

Notes

This paper represents the combined effort of all Veterans Administration Psychiatric Program Evaluation Staff personnel. The writers are particularly indebted to Lee Gurel and Robert Maroney, Program Evaluation Staff Director and Assistant Director respectively. Received for publication April 20, 1970.

1. Those continua for which discrimination appears to be based on an additive mechanism by which excitation is added to excitation at the physiological level (e.g., heaviness, loudness, brightness) are labeled prothetic. Metathetic continua are those for which discrimination behaves as though based on a substitution mechanism at the physiological level (e.g., pitch, position). See Stevens and Galanter, 1957.

2. While it was intended that the forms of the inventory used for category rating and for ratio estimation be entirely parallel, the format for the ratio estimations required Outcome 11 of the inventory to be split into two items. Since there could be some question about the equivalence of one or both of the split (ratio-estimation) items to the inventory (category-rating) item, all three items were omitted in the scale-comparison analyses.

3. The ϵ^2 indicates the proportion of dependent-variable variance accounted for by the independent variable. Since ϵ values are scaled like r values, we can follow a proposed convention for rs (Cohen, 1969) and consider the ϵ^2 values of .01, .09, and .25 to reflect small, medium, and large effects respectively.

4. A group's mean scores on the outcomes can be visualized as a profile, such as the MMPI.

References

Cohen, J. 1965. Some statistical issues in psychological research. In *Handbook of clinical psychology,* ed. B. Wolman, pp. 95–121. New York: McGraw-Hill.

Cohen, J. 1968. Multiple regression as a general data-analytic technique. *Psychological Bulletin* 70(6):426–443.

Cohen, J. 1969. *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Coombs, C. 1964. *Theory of data.* New York: Wiley.

Ekman, G., and Sjöberg, L. 1965. Scaling. In *Annual review of psychology,* ed. P. R. Farnsworth, pp. 451–471. Palo Alto, Calif.: Annual Reviews, Inc.

Galanter, E. 1962. Contemporary psychophysics. In *New directions in psychology,* introd. T. Newcomb, New York: Holt, Rinehart and Winston, pp. 89–156.

Lorei, T. W. 1970. Staff ratings of the relative importance of the consequences of release from or retention in a psychiatric hospital. *Journal of Consulting and Clinical Psychology* 34(1):48–55.

Lorei, T. W., and Cohen, J. 1970. Hospital and occupation group differences in opinion about the release or retention of psychiatric patients. *Journal of Clinical Psychology* 26(2):223–229.

Stevens, S. S. 1959. Measurement, psychophysics, and utility. In *Measurement: Definitions and theories,* ed. C. Churchman and R. Ratoosh, pp. 18-63. New York: Wiley.

Stevens, S. S. 1966. A metric for the social consensus. *Science* 151:530–540.

Stevens, S. S. 1968. Ratio scales of opinion. In *Handbook of measurement and assessment in behavioral sciences,* ed. D. K. Whitla, pp. 171–199. Reading, Mass.: Addison-Wesley.

Stevens, S. S., and Galanter, E. H. 1957. Ratio scales and category scales for a dozen perceptual continua. *Journal of Experimental Psychology* 54:377–409.

STRESS PRE-ENTRY IN MOTOR SYSTEMS

Donald G. MacKay
University of California, Los Angeles

Stress pre-entry is the occurrence, before it is due, of the stressed component in a series of rapidly produced movements. The phenomenon appeared in speech (experimentally produced spoonerisms) and in patterns of finger movement: serial-order errors were usually a stressed element entering before its time. The probability of serial-order errors was also found to increase as a function of rate of action. Several explanations are possible, but the model that best fits the data is one of a scanning device for determining serial order of rapidly produced behavior.

Errors in the serial order of action are crucial to theories of behavior in the same sense that illusions are crucial to theories of perception. But although the problem of serial order in behavior has attracted considerable theoretical attention (Milner, 1961; Conrad, 1965; Bryden, 1967; Wickelgren, 1969) since Lashley's famous paper (1961), few experiments on errors in serial order have appeared (MacNeilage, 1964; Shaffer and Hardwick, 1968, 1969). The experiments reported here focus on a limited aspect of the problem, namely, the effects of motor intensity (or stress) and rate of action on serial reversals in speech and in patterns of finger movement.

We began by examining the role of stress in spoonerisms (involuntary reversals in the serial order of phonemes in natural or conversational speech). Next we found a method of experimentally inducing spoonerisms that allowed a systematic test of the effects of stress and rate on serial reversals in speech. We then carried out an analogous experiment on serial-order errors in patterns of finger movement, our question being whether stress and rate play a similar role in motor systems other than speech. Finally we developed a model to account for the effects of stress and rate on serial-order reversals in these two motor systems.

SPOONERISMS IN GERMAN

Our interest in the effect of stress in serial-order errors began with an examination of 124 spoonerisms in German (Meringer and Mayer, 1895; Meringer, 1908). Our reasons for considering this corpus were, first, its size and, second, Meringer's care in collecting the data (see MacKay, 1970). It appeared that one of the reversed phonemes in these spoonerisms often originated in a stressed syllable and the other in an unstressed one. Moreover, the stressed element usually seemed to be the one that occurred before it was due, as in

vaga BUND → vabagund, or [1]
ge BRAUCH → begrauch, [2]

where the intruding B in both examples originated in a stressed syllable (capitalized; the arrow means 'was spoken as'). Our first question was whether this apparent effect of stress in spoonerisms exceeded chance expectation.

Our analyses were carried out within the following definitional framework. *Within-word spoonerisms* were defined as reversals of phonemes within a single word, as in Examples 1 and 2. Spoonerisms with reversed phonemes originating in different words were not included in this analysis.[1] An *intruding phoneme* was defined as the phoneme executed before its time, as B in Examples 1 and 2. A *lagging phoneme* was defined as the phoneme executed after its time, as G in Examples 1 and 2. The intruding and lagging phonemes originated in different syllables in all but one of the examples in Meringer's corpus. Since one syllable is more stressed than all the others in multisyllabic words in German,[2] *syllabic stress* was represented as 0 or 1, determined from a standard German dictionary (Wahrig, 1966). An instance of *stress pre-entry* was counted whenever the intruding pheneme in a spoonerism originated in a stressed syllable. Thus Example 3 represents an instance of stress pre-entry whereas Example 4 does not. Note that only one syllable was counted as stressed:

be GA bung → gebabung (example from Meringer) [3]
be GA bung → bugabeng (hypothetical example) [4]

Thus in each word, stress pre-entry implies that the lagging phoneme originated in an unstressed syllable. An instance of *stress post-entry* was counted when the intruding phoneme originated in an unstressed syllable and the lagging phoneme in a stressed one. Example 5 thus represents an instance of stress post-entry:

be GA bung → bebagung (hypothetical example) [5]

The last definition is of *corpus of natural speech:* this corpus, consisting of 230 correctly produced sentences from Meringer (comparable to those containing the spoonerisms), was used to determine the null hypotheses.

Two hypotheses guided our analyses. The *like-with-like hypothesis* was first proposed by Boomer and Laver (1968): "The origin syllable and the target syllable of a slip are metrically similar in that both are salient or both are weak." Using our definition of stress (or salience), this hypothesis predicts that both syllables in within-word spoonerisms should be unstressed with greater than chance probability. Moreover, stress pre-entry should be no more frequent than stress post-entry under this hypothesis. The *stress pre-entry hypothesis*, however, predicts that stress pre-entry should exceed chance expectation. That is, one of the reversed phonemes should be stressed with greater than chance probability, and stress pre-entry should be significantly more frequent than stress post-entry.

According to our analyses of Meringer's data, both syllables were unstressed in 18% of the German spoonerisms, whereas 26% would be expected by chance (based on the relative frequency of stressed and unstressed syllables in the corpus of natural speech). This difference was not statistically reliable at the .05 level using a chi-square test on the raw data.

However, stress pre-entry was significantly more frequent than chance expectation. Here chance expectation reflected the frequency with which stressed syllables followed unstressed syllables in the multisyllabic words in the corpus of natural speech—which was 23% of the time. But stress pre-entry appeared 71% of the time in the corpus of spoonerisms, a difference significant from chance at the .01 level, chi-square test. The same test showed that stress post-entry was significantly less frequent (11%) than chance expectation (51%, based on the frequency with which stressed syllables preceded unstressed ones in the multisyllabic words in the corpus of natural speech). Both findings strongly support the stress pre-entry hypothesis.[3]

SPOONERISMS IN ENGLISH

Is stress pre-entry a universal tendency common to all speakers regardless of language? As a preliminary test of this hypothesis, we analyzed an American corpus of 55 spoonerisms published by Bawden (1900; see MacKay, 1970), apparently without knowledge of Meringer's work. Stress was tabulated for within-word spoonerisms as before, with the results below. Since we lacked a corpus of correctly produced sentences from Bawden's speakers, we based the null hypothesis on a corpus of 255 multisyllabic words randomly sampled from the index at the top of odd-numbered pages of a large American dictionary.

According to our analysis of Bawden's data, stress pre-entry in the English spoonerisms (57%) was more common than chance expectation (11%, calculated as before); using a chi-square test, a difference significant at the .01 level. Similarly, stress post-entry was less common (28%) than chance expectation (55%), a difference significant at the .01 level, same test. Both findings support a stress pre-entry hypothesis for within-word spoonerisms in English.[4]

Both syllables were unstressed in the English spoonerisms only 15% of the time, whereas 34% would be expected by chance. Combining these data with those from the German spoonerisms, we found that both syllables were unstressed significantly less frequently than chance expectation at the .02 level, chi-square test. This finding directly contradicts the like-with-like hypothesis of Boomer and Laver (1968). The contradiction may be due to the fact that Boomer and Laver based their hypothesis on a conglomeration of errors that included substitutions, anticipations, perseverations, and omissions. But since the role of stress varies in different classes of errors (see MacKay, 1969), lumping different classes of errors together seems unwise. The like-with-like hypothesis simply does not hold for syllabic stress in within-word spoonerisms, although it may hold for other errors, given appropriate statistical analyses.

EXPERIMENTALLY INDUCED SPOONERISMS

We discovered that spoonerisms sometimes occur when subjects rapidly repeat a sequence of nonsense syllables such as 'TAY GAY BAY DAY.' In the experiment to be described, we systematically varied which syllable in this sequence was stressed. Under the stress pre-entry hypothesis we predicted that when spoonerisms did occur, one of the reversed components would be stressed with greater than chance expectation. Moreover, stress pre-entry in these experimentally induced spoonerisms should be more frequent than stress post-entry.

This experiment offered several advantages over our analyses of spoonerisms in natural speech. First, by using nonsense materials we could effectively rule out the semantic or 'deep' interpretations that have dominated the discussion of speech errors since Freud. Second, we could experimentally manipulate speech rate, a factor that went uncontrolled in observations of natural speech. Third, we could test the *informational* hypothesis of stress pre-entry. This hypothesis maintains that in natural speech, stressed syllables and words are more informative (in the information-theory sense) than unstressed ones, so that this may be why they enter before their time, and not because of motor intensity per se. By systematically varying which syllable was stressed we were able to determine whether motor intensity causes stress pre-entry independent of information value.

Method

–Subjects and apparatus–Twenty-two right-handed students at Edinburgh University were paid for their participation in the experiment. Their responses were recorded with a Revox tape recorder, and their errors were scored later from the tape.

–Experimental design–Two factors were varied, rate of speech and which syllable in the sequence was stressed.[5] Each subject was instructed to speak at two rates for each stress condition: maximum rate and normal rate. The order of the two rates was counterbalanced across subjects and stress conditions.

There were five stress conditions. In the no-stress condition, the subjects were instructed not to emphasize any of the syllables. In the remaining four conditions, one of the syllables was emphasized (made louder than the others), giving Examples 6 to 9, stressed syllables capitalized. This particular sequence

TAY gay bay day [6]
tay GAY bay day [7]
tay gay BAY day [8]
tay gay bay DAY [9]

of syllables was chosen to maximize the probability of serial-order errors.[6] But note that our hypothesis depends on whether an element in the sequence was stressed or unstressed, a systematically varied factor that was independent of the particular sequence chosen.

In a practice session before the experiment, the subject repeated each stress pattern five times at both rates of speech. In the experiment proper, he was instructed to continue repeating the sequence until told to stop. For both rate conditions, this stop signal came after 20 repetitions of the sequence. Order of the stress conditions was randomized across subjects.

Errors were scored in a pattern only when the first syllable was produced correctly.

Results

Spoonerisms were defined as before. For example, 'TAY bay gay day' represents a spoonerism, whereas 'TAY bay bay day' does not. The probability of spoonerism per sequence of four syllables was .015 (averaged over stress and rate conditions). As might be expected, the maximal rate (.280 sec per syllable) was considerably faster than the normal rate (.358 sec per syllable). As to frequency of spoonerisms for the two rates, 57% of the spoonerisms occurred at maximal rate and 43% at normal rate, a difference reliable at the .05 level, chi-square test. The importance of this effect of rate for theories of spoonerisms is outlined in the discussion.

Of the spoonerisms 95% occurred between adjacent syllables in the sequence. Moreover, one of the reversed components was stressed much more often than chance expectation. Only 40% of the reversals should involve a stressed component by chance, where chance reflects the frequency of stressed and unstressed elements in all of the experimental conditions. But in fact, 51% of the reversals involved a stressed component (see Table 1), which exceeds chance expectation at the .05 level, chi-square test.

Table 1. Experimentally induced spoonerisms as a function of stress

	One syllable stressed	Neither syllable stressed	Stress pre-entry	Stress post-entry
Reversals (percent)	51	49	90	10
Chance (percent)	40	60	50	50

Moreover, the experimental spoonerisms obeyed a stress pre-entry rule: the stressed element entered before it was due much more frequently than after. These data are shown in Table 1, where it can be seen that stress pre-entry was much more frequent than chance expectation, and stress post-entry much less frequent than chance, a statistically reliable finding at the .01 level, chi-square test.

SERIAL ORDER IN PATTERNS OF FINGER MOVEMENT

Is stress pre-entry a universal error tendency common to all motor systems or does it only appear in speech? As a first step in answering this question, we examined serial-order errors in patterns of finger movement. The experiment was analogous to that reported above for speech: the subjects rapidly repeated a sequence of finger taps on telegraph keys. As before, the stress pre-entry hypothesis was tested, here by systematically varying which finger tap was emphasized. Our prediction was that, as in speech, stressed movements would occur before their time more frequently than chance expectation.

Method

-Subjects-The subjects were the same 22 students who had participated in the experiment just reported, the order of the two experiments being counterbalanced across subjects.

-Apparatus-The main piece of apparatus was a four-channel Gerbrands event recorder, which recorded the time and duration of the responses. A response was a 1-mm depression of a standard telegraph key. There were four keys in all, one for each finger of the left hand. By placing his left arm on the table, the subject could easily rest each finger on one of the keys, which were mounted on a piece of bristol board attached to the table.

-Experimental design-The movement pattern in all experimental conditions was exactly the same. If the fingers of the left hand are labeled *a* to *d* from little finger to index finger, then the pattern was *acbd*. It might be argued that this particular sequence encourages the serial error *abcd*. This may be true, but our question was whether the *abcd* errors became more frequent when *b* was stressed than when it was not.

Again, the independent variables were the rate of executing the sequence and the stress pattern. For each stress condition the subject ran through the sequence at two rates: as fast as he could, and at a normal rate. Order of the two rates was counterbalanced across subjects and stress conditions. As before, there were five stress conditions: a different finger was stressed or emphasized in four of the stress conditions, and in the fifth, none of the fingers was stressed. Order of the stress conditions was randomized across subjects. In the actual experiment, the subject repeated each movement pattern 20 times for each rate and stress condition (following a practice session in which he produced each stress pattern five times at both rates).

As before, errors were scored only when the first element was produced correctly.

Results

The probability of reversal was .01 per movement pattern, averaged over stress and rate conditions. As might be expected, the actual rate of movement (in seconds per key press) diminished with maximal-rate instructions: the response rate was .119 for the slow rate and .088 for the maximal rate. And more serial-order errors occurred at the maximal rate (60%) than at the slow rate (40%), a finding significant at the .05 level, chi-square test. For some reason, serial reversals became more likely with faster rates of movement.

Again, the actions reversed were usually temporally adjacent in the sequence, and they frequently involved a stressed component. As can be seen in Table 2, one of the movements was stressed in 53% of the reversals, significantly more than the 40% expected by chance (.05 level, chi-square test). Also shown in Table 2 is the probability of stress pre-entry and post-entry (defined as for speech). Stress pre-entry was significantly more frequently than chance expectation at the .01 level, chi-square test. Similarly, stress post-entry was less frequent than chance expectation (.01 level, same test). As in spoonerisms, errors in the serial order of finger movements obeyed a law of prior entry, with stressed elements occurring before their time.

Table 2. Serial-order errors in patterns of finger movements as a function of stress

	One tap stressed	Both taps unstressed	Stress pre-entry	Stress post-entry
Reversals (percent)	53	47	86	14
Chance (percent)	40	60	50	50

-Controls-Two of the preceding conditions were rerun on 11 right-handed students at UCLA. Conditions *acBd* (stress on finger *B*) and *acbd* (no stress) were produced 20 times at maximal rate (with practice trials as before). Then four control conditions were added to test the hypotheses outlined below, the order of all conditions being counterbalanced across subjects.

The *von Restorff hypothesis* was one of those tested. Since only one element was stressed in the critical conditions of the main experiment, we wondered whether this element, being different from the others, stood out and was therefore more prone to serial-order reversal. To test this, we ran a condition where all the finger taps but one were stressed. The finger pattern was *ACbD*, stressed fingers capitalized. Under the von Restorff hypothesis we predicted that the unstressed element in *ACbD* would occur before its time as often as the stressed element in *acBd*. However, the data showed that stress pre-entry in *acBd* was much more common than unstressed pre-entry in *ACbD*. Finger *b* occurred before its time more frequently when stressed than when unstressed, a difference significant at the .05 level, chi-square test. This finding conclusively rules out an explanation of stress pre-entry as a von Restorff effect.

The second hypothesis tested in the control conditions was the *relational hypothesis*. The reason for stress pre-entry may lie in the relation between stressed and unstressed elements, the stressed elements entering before unstressed elements but not before other stressed elements. To test this, we ran a control condition in which all of the elements were stressed. The data showed that serial-order reversals were no more common when all of the elements were stressed ($N = 5$) than when all were unstressed ($N = 5$), which supports the relational hypothesis. Moreover pre-entry of finger *b* was more frequent in condition *acBd* than in condition *ACBD*, although this difference just missed statistical reliability at the .05 level using a chi-square test. Nevertheless, these data tentatively support the relational hypothesis, since stressed elements tended to advance on unstressed elements but not on other stressed ones.

The last hypothesis tested here was the *specificity hypothesis*. In order to rule out the possibility that our results were specific to the left hand, we ran condition *acBd* using the right hand of our right-handed subjects. Serial-order errors were much less frequent with the right hand, suggesting a dexterity effect. Nevertheless a stress pre-entry trend was observed in the right hand, although it failed statistical significance because of the small number of errors.

Finally, the pattern *dbCA* was run with the left hand to test the generality of the stress pre-entry phenomenon. Stress pre-entry occurred with about the same frequency in *dbCa* as in *acBd*, indicating that stress pre-entry is not tied to a particular movement pattern.

A MODEL OF STRESS PRE-ENTRY

Serial-order reversals in sensory or input systems bear a curious resemblance to stress pre-entry in motor or output systems. When two visual stimuli of similar form are presented in succession, a dim one first and a very intense one second, the second is sometimes perceived as preceding the first. The probability of this reversal in the serial order of perception can be expressed as

$$P(R) = K(b - a)\,, \qquad [10]$$

where $P(R)$ is the probability of reversal, K is a constant, and a and b are the intensities of the first and second stimuli respectively (Korte, 1915). In this last section, we explore the possibility that motor reversals can be explained with an analogous formula.

There are three rather well documented premises that can be used to explain stress pre-entry in sensory systems:

Premise 1: Excitability of the set of analyzers for a form must be boosted to threshold before the form is perceived (after Sutherland, 1959, and others).

Premise 2: A strong or intense stimulus boosts these analyzers to threshold faster than a weak one. This assumption is backed by both psychophysical and physiological evidence (Pulfrich, 1922; Liang and Piéron, 1942; Lit, 1949; Wilson and Anstis, 1969; Adrian and Matthews, 1927).

Premise 3: The mechanism for judging the serial order of input is independent of the analyzers themselves, in the sense that the rate of activation of the analyzers is not taken into consideration in judgments of serial order.

Although these premises do not encompass all aspects of perception, they suffice to explain perceptual reversals, since a strong stimulus reaching analyzer threshold before a weak one is judged to precede the weak one, even though the weak one preceded in the external world.

The analogous premises for explaining stress pre-entry in motor systems might then take the form:

Premise 1A: Excitability of a set of motor units for an action must be boosted to threshold for the action to be produced.

Premise 2A: Motor units for stressed actions are boosted to threshold faster than those for unstressed actions.

Premise 3A: The mechanism for sequencing behavior is independent of the motor units themselves (as Lashley, 1961, suggested).

Two objections to this analogy between sensory and motor reversals might be raised. One is that perceptual reversals only occur when visual stimuli are separated by 100 msec or less (Wilson and Anstis, 1969), whereas more than 100 msec usually separate the reversed components in speech. However, it is highly likely that motor reversals originate at a central rather than peripheral level—for example, in higher-level instructions to the motor units rather than in the motor units themselves. If the rate at which these instructions are given at this central level exceeds the speed at which the final output is produced, the objection is overcome. The objection also becomes irrelevant if time constants in motor units are longer than in perceptual analyzers even though similar principles operate.

A second objection to the analogy might be that stressed output differs from unstressed output in other ways besides intensity. Stressed actions last longer than unstressed ones, and they involve faster rates of movement. However, these concomitants of stress can be predicted, as we will see as we develop our premises into a general model for motor systems.

To develop that model, let us first consider the possibility that sequences of acts to be rapidly performed are represented or displayed in abstract form and proper order in an output buffer, such that the elements in the buffer prime or partially activate corresponding motor units. Independent justification for the assumption of priming or subthreshold activation of motor units prior to final motor output is found in such other errors in speech as anticipations (Lashley, 1961). If this priming varies in direct proportion to stress, then stressed units are closer to threshold (prior to final output) than are unstressed units.[7] Some justification for this assumption is found in omissions of speech sounds: the sounds omitted are usually unstressed, rather than stressed (Heffner, 1964).

Let us next consider the possibility that a scanning mechanism determines the serial order of behavior by sweeping over the buffer system in undirectional fashion. By boosting the excitability of the buffer elements, this scanning device eventually triggers the motor units, but at the same time is unconnected to or independent of the motor units themselves (Premise 3A). The scanner, it is proposed, provides a linear boost in excitability of elements in the buffer for a period determined by the rate of scanning. Voluntary control of the rate of scanning varies the rate of action within limits, the upper limit being determined by the duration of temporal summation needed to activate the motor units. If stressed units are already closer to threshold than unstressed units, this summation time is less for boosting stressed units to threshold than unstressed ones (Premise 2A). This postulate allows a precise formulation of the probability of stress pre-entry [$P(S)$]. If A and B are units in serial order in the buffer system with summation times a and b respectively (i.e., a and b represent the time following onset of scanning needed to boost the motor units for A and B to threshold), then

$$P(S) = K(a - b - c)\,, \qquad [11]$$

where K is a constant and c is the time between onset of scanning A and B. Consequently,

$$\begin{aligned} P(S) = 0 \quad &\text{if} \quad a - b - c > 0 \\ &\text{or} \quad a < b + c\,, \quad \text{and} \qquad [12] \end{aligned}$$

$$\begin{aligned} P(S) > 0 \quad &\text{if} \quad a - b - c < 0 \\ &\text{or}\; a > b + c\,. \qquad [13] \end{aligned}$$

These formulas derived from the model can be seen to summarize all of our findings. Note that $P(S)$ increases as c becomes smaller. One of the factors influencing c is the separation of the relevant units in the buffer system. All other factors being equal, serial-order reversals between elements in close proximity are to be expected. Thus the fact that reversed phonemes occur closer together in natural speech than would be expected by chance (see MacKay, 1970; Cohen, 1967) is represented in these formulas. Similarly, the fact that reversed actions were usually temporally adjacent in our experiments is captured in Formula 13.

Another factor that determines c (the time between onset of scanning the relevant units in the buffer) is the rate of scanning or, equivalently, the rate of action. With faster rates of scanning, c becomes smaller and $P(S)$ becomes greater in Formula 13, which captures the fact that serial-order reversals became more likely with rapid rates of speech and finger movement in our experiments. Finally, these formulas summarize both the stress pre-entry phenomenon itself and the relational hypothesis by predicting a high frequency of serial-order reversals when a is large and b is small—that is, when the first element is unstressed and the second is stressed, as is usually the case in motor reversals.

Thus developed, the model also explains why stressed speech sounds are longer than unstressed ones (Heffner, 1964). If the motor units for stressed segments reach threshold sooner than those for unstressed ones, then, all other factors being equal, stressed sounds in the final output are bound to be longer than unstressed ones. And finally, the model explains why stressed actions are carried out with faster rates of action than unstressed ones (O'Hala, 1969). The rate of muscular contraction is determined by the number of impulses per second from the motor units (Buller, Eccles, and Eccles, 1960). If the activity of the motor units prior to actual movement is greater for stressed than unstressed actions, then following scanning in this model the final activity of stressed units is also greater, leading to faster rates of movement.

Several limitations of the model should be emphasized. First, it applies to only one aspect of serial-order reversals; many other factors known to influence errors in serial order have been ignored for the sake of simplicity. Second, it is based on a restricted type of behavior; we have no grounds for generalizing our model to the natural behavior patterns and rapidly executed thought sequences that originally sparked Freud's interest in faulty actions. Perhaps Freud was partially right in contending that for natural speech "even the apparently simple cases of speech blunder will be traced to a disturbance caused by a half-repressed idea outside of the intended context" (1914, p. 66). However, Wundt's hypothesis (1907) that these ongoing mental activities often act as distractors (i.e., as catalysts rather than causes of errors) warrants serious investigation as an alternative hypothesis (see n. 5).

Other theories

Some aspects of scanning models are necessarily vague at this point, partly because the nature of the output buffer is as yet unknown. The scanning mechanism developed in the model above is essentially analogous to certain features of a television receiver. But there could be a different scanning mechanism, one based on the principle of the tape-recorder head, which 'reads out' the electromagnetic pattern on a tape. Instead of energizing elements in an interim storage system and thereby activating the motor components, this scanner would serially read out from the buffer a set of values for the motor units. This 'read out' scanner can be made to predict the same findings as the 'receiver' scanner in the model above, and at present I see no way of experimentally distinguishing the two.

Another possibility is that the scanner may release from inhibition the programs in the output buffer rather than activate them directly. This view is related to Lashley's hypothesis that "it is as if [an] aggregate of words were in a state of partial excitation, held in check by the requirements of grammatical structure, but ready to activate the final common path, if the effectiveness of this check is in any way interfered with" (1961, p. 186). But again, viewing the scanner as a device for serial disinhibition leads to the same predictions as the scanning mechanisms discussed just above: our data tolerate all three versions of the scanner.

Even so, the phenomenon of stress pre-entry presents major problems for other theories of serial order. For example, Conrad's (1965) proposal that serial-order errors are multiple substitutions occurring simply by chance provides no insight into stress pre-entry and is essentially the null hypothesis that our data refute.

Wickelgren's (1969) chain-association model of spoonerisms must also undergo serious revision to explain stress pre-entry. Chain association represents a radical departure from the mechanisms considered so far. In associative models, the motor units, rather than being independent of the serial-ordering mechanism, are linked by directional bonds, so that an entire word may be released in order by triggering the first motor unit. Associative models have difficulty explaining how sequences are continued after a serial-order reversal occurs. That is, if the correct sequence is *ABCD* and the first two elements are reversed, present chain-association models predict the continuation *BABCD* rather than the *BACD* that we find. This difficulty may be overcome if the links serve to boost the activation of the motor units to threshold and these units are viewed as variable resistors in the chain rather than all-or-none devices. Moreover, the only way I can see for associative models to handle stress pre-entry is to represent stress in the links between motor elements, rather than in the motor elements themselves. Further research to test these new assumptions is needed before chain association can be considered a viable principle in motor systems.

This résumé of alternative theories is by no means comprehensive but should perhaps suffice to refute Freud's assertion that "no matter what may be said about faulty actions, the whole subject is known to everybody as something self-evident" (1914, p. 114).

Notes

This work was supported by UCLA Grant 2426 and USPHS Grant 16668-01. The author thanks T. Falbo for running subjects and analyzing data, Professor R. C. Oldfield for providing space and equipment for the study at the M. R. C. Speech and Communication Unit, Edinburgh University, and Dr. G. Soderberg for running the control conditions. Received for publication June 19, 1970.

1. This was because analysis of stress, defined in terms of relative motor intensity, could be reliably determined only for syllables within a single word.

2. Comparison of syllabic stress in different words in a written corpus was out of the question, since different words may there be stressed to different extents. This 'word stress' would alter the relative motor intensity of syllables in different words but not within a single word, and could only be determined from a tape recording of the error. Thus, analysis of stress in our corpus seemed justified only for within-word spoonerisms, not for between-word spoonerisms.

3. Analysis of between-word spoonerisms showed a slightly different pattern of stress. Of the between-word reversals in Meringer's corpus 97% occurred in syllables that were both stressed, an outcome exceeding chance expectation at the .01 level, chi-square test. But this finding in no way opposes a stress pre-entry hypothesis, since the second word may have received greater stress than the first as predicted in the hypothesis. We had no way of testing this question with our present data. Nevertheless, Boomer and Laver were able to analyze the role of word stress in a corpus of tape-recorded spoonerisms, and their analysis seems to support the stress pre-entry hypothesis: the second word in between-word spoonerisms was usually emphasized in their data, suggesting that "the tonic syllable is the only element with sufficient neurophysiological prominence to break through the inhibitory forces holding the programme in check" (1968).

4. Analysis of the between-word reversals in English showed the same pattern of stress as in German. In 88% of the corpus the syllables containing reversed phonemes were both stressed. Again, however, we had no way of analyzing the word stress superimposed on the syllabic stress, so that these data may also support the stress pre-entry hypothesis (see n. 3).

5. An additional variable was manipulated in this and the next experiment. Half the subjects heard a tape recording of Strauss's "Thunder and Lightning Polka" while producing the sequences. This tended to increase the probability of all error types, but its detailed effect is left for a later report, since the effect on serial-order errors alone failed statistical significance.

6. That is, consonant reversals are most likely when similar consonants are preceded or followed by identical vowels (see MacKay, 1970).

7. Note that this view of stress would also explain why stressed syllables predominate in the partial recall of words (see Brown and McNeil, 1965).

References

Adrian, E. D., and Matthews, R. 1927. The action of light on the eye. *Journal of Physiology* 63:378–414.

Bawden, H. H. 1900. A study of lapses. *Psychological Monographs* 3(4):1–121.

Boomer, D. S., and Laver, J. D. M. H. 1968. Slips of the tongue. *British Journal of Disorders of Communication* 3(1):1–12.

Brown, R., and McNeil, D. 1965. The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior* 3:325–327.

Bryden, M. P. 1967. A model for the sequential organization of behavior. *Canadian Journal of Psychology* 21(1):37–56.

Buller, A. J., Eccles, J. C., and Eccles, R. M. 1960. Interaction between motorneurons and muscles in respect of the characteristic speeds of their responses. *Journal of Physiology* 150:417–439.

Cohen, A. 1967. Errors of speech and their implication for understanding the strategy of speech users. *Zeitschrift für Phonetik* 20:177–181.

Conrad, R. 1965. Order errors in immediate recall of sequences. *Journal of Verbal Learning and Verbal Behavior* 4:161–169.

Freud, S. 1914. *Psychopathology of everyday life,* trans. A. Brill. New York: Penguin.

Heffner, R. M. S. 1964. *General phonetics.* Madison: University of Wisconsin Press.

Korte, A. 1915. Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie* 72:123–126.

Lashley, K. S. 1961. The problem of serial order in behavior. In *Psycholinguistics,* ed. S. Saporta, pp. 180–198. New York: Holt, Rinehart and Winston.

Liang, T., and Piéron, H. 1942. Récherches sur la latence de la sensation lumineuse par la méthode de l'effet chromostéréoscopique. *Année Psychologique* 1:1–53.

Lit, A. 1949. The magnitude of the Pulfrich stereophenomenon as a function of binocular differences of intensity at various levels of illumination. *American Journal of Psychology* 62:159–181.

MacKay, D. G. 1969. Forward and backward masking in motor systems. *Kybernetik* 2:57–64.

MacKay, D. G. 1970. Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia* 8:323–350.

MacNeilage, P. F. 1964. Typing errors as clues to serial ordering mechanisms in language behavior. *Language and Speech* 7:155–159.

Meringer, R. 1908. *Aus dem Leber der Sprache: Versprechen, Kindersprache, Nachahmungstrieb*. Berlin: Behr's Verlag.

Meringer, R., and Mayer K. 1895. *Versprechen und Verlesen: Eine psychologisch-linguistische Studie*. Stuttgart: Göschensche Verlagsbuchhandlung.

Milner, P. M. 1961. A neural mechanism for the immediate recall of sequences. *Kybernetik* 1:76–81.

O'Hala, J. J. 1969. Aspects of the control and production of speech. Unpublished doctoral dissertation, University of California, Los Angeles.

Pulfrich, G. 1922. Die Stereoskopie im Dienste der isochromem und heterochromen Photometrie. *Naturwissenschaften* 10:533–761.

Shaffer, L. H., and Hardwick, J. 1968. Typing performance as a function of text. *Quarterly Journal of Experimental Psychology* 20:360–369.

Shaffer, L. H., and Hardwick, J. 1969. Errors and error detecting in typing. *Quarterly Journal of Experimental Psychology* 21:209–213.

Sutherland, N. S. 1959. Stimulus analyzing mechanisms. In *Proceedings of the symposium on the mechanization of thought processes*, vol. 2, pp. 575–609. London: H. M. S. O.

Wahrig, G., ed. 1966. *Das grosse deutsche Wörterbuch*. Munich: Bertelsmann Verlag.

Wickelgren, W. A. 1969. Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review* 76:1–15.

Wilson, J. A., and Anstis, S. M. 1969. Visual delay as a function of luminance. *American Journal of Psychology* 82:350–358.

Wundt, W. 1907. *Outlines of psychology*, trans. C. H. Judd. Leipzig: Wilhelm Engelmann.

EFFORT IN SUSTAINED AND PHASIC HANDGRIP CONTRACTIONS

William S. Cain and Joseph C. Stevens
John B. Pierce Foundation Laboratory and Yale University

Subjects maintained handgrip contractions over time at constant levels of perceived effort. The resulting functions relating force to time were similar in overall shape from person to person and highly reliable for any given person. The reduction in exerted force over time was well described by an equation that comprises two exponential terms; the rate parameters of both terms are independent of the initial force of contraction. Prior fatigue, as generated by sustained constant-effort contractions, increased the perceived intensity of phasic contractions. The psychophysical power law described the perceived intensity of phasic contractions both before and after the exertion of fatiguing contractions; fatigue had little effect on the size of the exponent.

The degree of perceived effort that accompanies an isometric contraction depends mainly on how much force is exerted and for how long. When a subject maintains a contraction at a constant level of force, he reports a continuous increase in his subjective effort over time. Accordingly, when instructed to maintain a prolonged contraction at a constant level of subjective effort, he continuously reduces the force of the contraction, through at least the first few minutes. These two tasks—maintenance of constant force and maintenance of constant effort—represent complementary ways to study the course of proprioceptive sensations associated with fatigue.

With respect to the first of these tasks (constant force), Stevens and Cain (1970) have found that the degree of apparent effort in a handgrip contraction maintained at a constant force level grows as a power function of duration. The rate of growth (i.e., the exponent) of the power function is approximately the same for any force of contraction. The size of the exponent is about .6, meaning that effort increases approximately fourfold for any tenfold increase in duration, irrespective of the physical load. This relation between apparent effort and duration provides yet another instance of the psychophysical power law shown by S. S. Stevens (1960) to apply with great generality throughout the sensory domain.

With respect to the second task (constant effort), Eason earlier (1959) reported that in order to maintain the same effort in an isometric elbow flexion, subjects decreased the force level (from the initial load of 30 lb) approximately as an exponential function of time. The two experiments reported below extend the study of handgrip to contractions maintained over time at several different constant subjective levels. The purpose of Experiment I was to determine whether the constant-effort function for handgrip is exponential (or something else) at all levels and to discover how the constants of the functions might depend on the initial level of force. The purpose of Experiment II was to learn how the fatigue induced by constant-effort contractions affects the perceived intensity of subsequent phasic contractions.

EXPERIMENT I: NATURE OF THE CONSTANT-EFFORT FUNCTIONS

Method

-Apparatus–The handgrip dynamometer was machined from a single heavy piece of tempered aluminum. It was shaped much like a hairpin, with its two arms parallel to each other. The arms were 22 cm long. Compression applied across them was converted into an electrical signal by capacitor plates mounted opposite each other on the inside surfaces of the arms. The plates formed part of an LC network. Any small movement of the arms relative to each other changed the areal overlap of the plates, and thereby the capacitance. The capacitance constituted one leg of an a.c. bridge, whose rectified output was fed to a d.c. recorder. The relation between d.c. output voltage and force applied to the dynamometer was found to be stable. A squeeze of 445 newtons (100 lb) produced a 3-mm excursion of the arms at the particular point where they were to be grasped by the subject.

The output from the bridge was also fed to a d.c. meter. A single reference mark on the dial of the meter served as a monitor for the subject. Each trial began when the subject exerted the force necessary to align the pointer of the meter with the reference mark. An offset potentiometer connected to the meter allowed the experimenter to vary the force necessary for alignment.

The testing took place in a sound-treated room. The subject sat in a chair of adjustable height that enabled him to adopt a comfortable position for grasping the dynamometer with the preferred hand.

-Procedure-Twelve male subjects participated (mean age, 21). Each test session had two parts. The first called for 15 phasic contractions, the second for 7 sustained (constant-effort) contractions. The first part used the method of magnitude estimation without a designated modulus; this task called upon the subject to match numbers to the apparent intensity of contractions of seven levels of force ranging from 88 to 429 newtons (19.8 to 96.4 lb) in equal logarithmic steps. The force level in the first trial was chosen from near the middle of the stimulus range; the subject was asked to assign to its apparent intensity any number he felt appropriate. To subsequent contractions, he was asked to give numbers in proportion to apparent intensity. Each of the seven force levels was exerted twice in the session in irregular order. (The numerical assignment for the first contraction was not counted.)

In the second part of the test session, the subject maintained one constant-effort contraction for each of the seven force levels previously judged in the first part of the session. A trial began when the subject aligned the indicator with the reference mark of the d.c. meter. When alignment was accomplished (approximately 2 sec), the meter was immediately concealed and the subject thereafter attempted to keep perceived effort constant for 2.5 min.

Results

Figure 1 shows the geometric means of the magnitude estimations of the phasic contractions plotted as a function of force level. The straight line fitted to the data in log-log coordinates is a power function with a slope (exponent) of 1.5. This slope is close to those found in similar studies by Stevens and Mack (1959) and Stevens and Cain (1970) when the task was the magnitude estimation of phasic contractions. However, the results from several investigations employing a variety of psychophysical methods suggest that 1.7 is a better estimate of the exponent (for a summary, see Stevens and Cain, 1970).

Figure 2 shows how force level was reduced over time to keep perceived effort constant. Even contractions of moderate initial intensity had to be sharply reduced during the first few seconds to prevent buildup of effort. The higher the perceived effort of the contraction, the greater the need to reduce the force, especially in the beginning; the functions therefore tend to converge rapidly at first, less rapidly later on.

The group data in Figure 2 reflect quite faithfully the constant-effort contractions for individual subjects. Although there were undoubtedly differences in strength and motor ability among the subjects, there was a pronounced uniformity in overall performance. Table 1 gives the range of variability of the group results at each

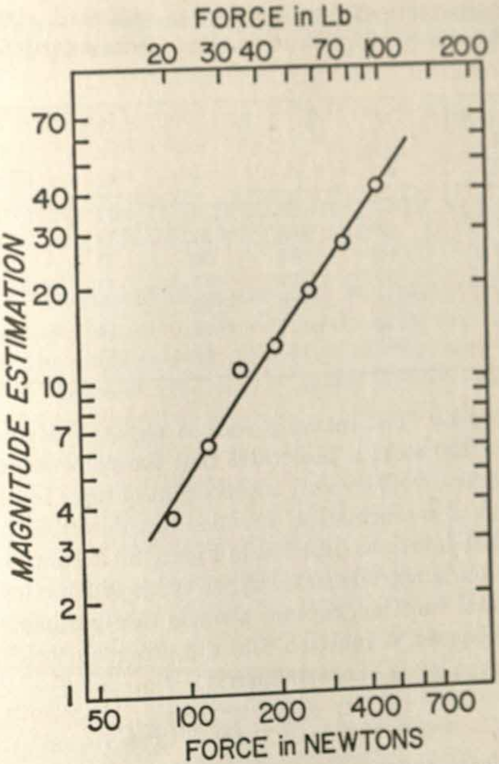


Fig. 1. Perceived intensity (geometric means) as a function of force for phasic contractions

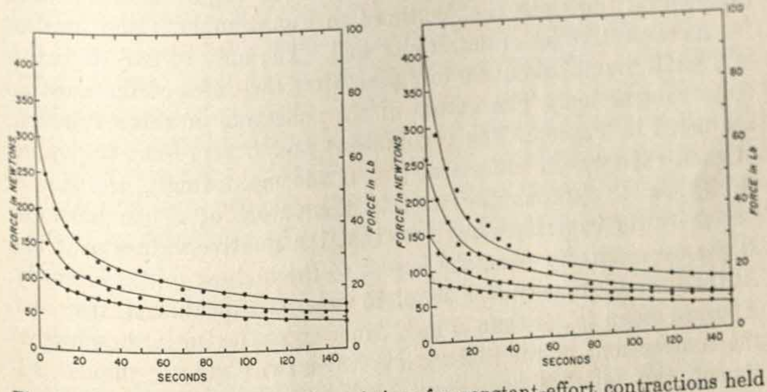


Fig. 2. Median force as a function of time for constant-effort contractions held at seven different levels of effort

Table 1. Parameters of the double-exponential function ($F = Ae^{-k_1 t} + Be^{-k_2 t} + C$) fitted to the constant-effort contractions and the range Q_3/Q_1 for each level of contraction

Initial force level: in newtons (preset)	A	B	C	k_1	k_2	Range of Q_3/Q_1
1: 88	4	60	22	.11	.01	1.1–1.7
2: 116	25	65	24	.11	.01	1.2–1.6
3: 150	60	59	36	.11	.01	1.2–1.4
4: 195	87	77	33	.11	.01	1.3–2.0
5: 254	111	114	22	.11	.01	1.2–1.6
6: 331	170	117	36	.11	.01	1.1–1.4
7: 429	267	105	49	.11	.01	1.4–1.7

level of intensity. The measure used is Q_3/Q_1 (calculated from $t = 10$ to $t = 150$ sec); it illustrates that the percentage variability in the contractions was roughly independent of force level.

What kind of mathematical function can describe the family of constant-effort functions depicted in Figure 2? A simple exponential decay function is not adequate. Other types of functions that, like the exponential function, contain a single rate-parameter were also deemed inadequate. A function that did provide a faithful description of all seven levels of contraction is

$$F = Ae^{-k_1 t} + Be^{-k_2 t} + C\,, \qquad [1]$$

where F is force and t is time. This equation is simply the sum of two exponential terms, each with its respective rate of decay.

The smooth curves fitted to the data in Figure 2 are double-exponential functions (i.e., conform to Equation 1). Each function fits its respective set of data quite well. (The only serious departure is a small 'hump' about 20 to 30 sec after the onset of the most intense contraction.) The values of the constants of these functions are listed in Table 1. The constants A and B represent the values of each exponential process at $t = 0$ and, accordingly, the sum of A, B, and C approximates closely the values of F (in newtons) preset by the experimenter. Note that the relative values of A and B change markedly from the lowest to the highest initial intensity. At the lowest intensity B is about 15 times greater than A. However, A grows much faster than B as a function of the initial intensity of the contraction; at the highest level A is twice greater than B. Although the relative contribution of each exponential component changes strikingly from level to level, the rate parameters k_1 and k_2

do not. All of the functions in Figure 2 have the same rates of decay. The effect of k_1 is most apparent in the beginning of each contraction, where very rapid decreases in force level were necessary to maintain effort constant. The effect of k_2 becomes most apparent in the later stages.

EXPERIMENT II: THE EFFECT OF PRIOR FATIGUE ON PHASIC CONTRACTIONS

The aim of Experiment II was to learn how prior fatigue affects the perceived intensity of phasic contractions. In the more purely receptive sense modalities (e.g., vision), there has been interest in how the relation between perceived intensity and physical intensity is affected by adaptation. The greater the degree of light adaptation, the higher the exponent of the psychophysical power function for brightness (Stevens and Stevens, 1963). Adaptation in olfaction (Cain and Engen, 1969) and taste (McBurney, 1966; Bartoshuk, 1968) also increases the exponent. Since muscle fatigue may be thought of as an analogue to sensory adaptation in that both phenomena result from prolonged continuous stimulation, it was of interest to know whether fatigue also affects the rate of growth of the psychophysical function for phasic contractions.

Each constant-effort function obtained in Experiment I provides a measure of the course of fatigue for a contraction maintained at a single level of apparent effort. The functions might also provide a means to relate the degree of fatigue produced by a contraction maintained at one level to the degree of fatigue produced by a contraction maintained at another level of effort. One possibility is that two different constant-effort contractions may arouse the same degree of fatigue when the areas under their respective force versus duration functions (Figure 2) are equal—that is, when the contractions are equal in terms of force level integrated over time. This possibility was explored in Experiment II, in which constant-effort contractions were used to generate muscular fatigue that was then assessed by its effect on the psychophysical function for phasic contractions.

Method

-Apparatus-The dynamometer described in Experiment I was used.

-Procedure-Nineteen males served as subjects (mean age, 22). Each session began with a series of phasic contractions. These were included merely to give the subject practice in making numerical estimates of intensity before the actual test trials began. Then came 14 test trials, 2 at each of the seven levels of force used in Experiment I. Half of the trials (the even-numbered ones) were 'fatigue trials' and half (the odd-numbered ones) were 'nonfatigue trials.' On a fatigue trial, the subject maintained a constant-effort contraction (i.e., a fatiguing contraction) for a specified duration and then immediately exerted a phasic contraction at one of the seven test levels. He assessed the apparent intensity of the phasic contraction (the test contraction) by magnitude estimation. On a nonfatigue trial, the subject exerted two phasic contractions separated by 8 sec, and as on the fatigue trials, only the second contraction was judged.

The initial force level of the fatiguing contraction was constant throughout a session, and the force level of the first phasic contraction on each nonfatigue trial was always the same as the initial force level of the fatiguing contraction. Therefore, each of the 14 test contractions within a session was preceded by a phasic or sustained contraction at one given level of apparent intensity. Pilot experiments had shown that this procedure allowed the numerical estimates obtained on fatigue trials to be most meaningfully compared to those obtained on nonfatigue trials. The interval between a nonfatigue and the following fatigue trial was 30 sec, and the interval between a fatigue and the following nonfatigue trial was 5 min.

There were six conditions of fatigue; four comprised Group I and two comprised Group II. The initial force levels and durations of the four fatiguing contractions in Group I were chosen to represent approximately equal areas (4,500 N.-sec) under four of the curves shown in Figure 2. That is, they were chosen to be equal in terms of force level integrated over time. Their initial force levels and their durations were: 88 N., 72 sec; 116 N., 62 sec; 195 N., 37 sec; 331 N., 21 sec. The two fatiguing contractions in Group II were also chosen to be equal in terms of force level integrated over time, but represented approximately twice the area (9,000 N.-sec) under the curves in Figure 2 as the contractions in Group I. Their initial force levels and durations were 116 N., 168 sec; and 331 N., 60 sec. If the degree of fatigue aroused reflects integrated force level, there should be no difference among the effects of the four fatiguing contractions in Group I and between the effects of the two fatiguing contractions in Group II.

Results

There was noteworthy stability in the constant-effort (fatiguing) contractions for individual subjects. Each of the eight sets of records in Figure 3 shows the performance of an individual subject on five contractions exerted in the same test session. The sets of records displayed were chosen at random for the 19 sets obtained for each fatiguing contraction from Group II, but they reflect the consistency of performance that characterized the entire experiment. There was a clear similarity in the overall course of the contractions from

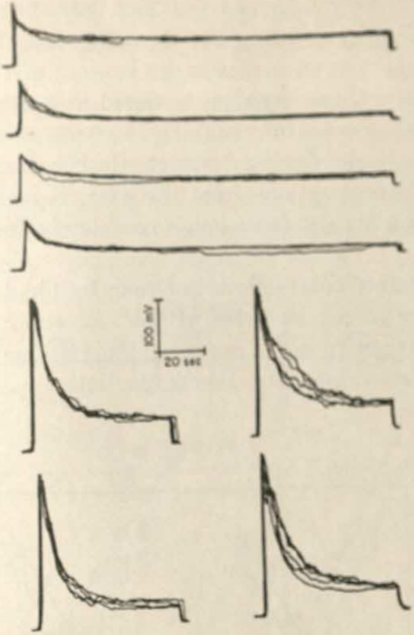


Fig. 3. Eight sets of records that show the performance of individual subjects on five repeated constant-effort contractions

one subject to another, but there was also clear variation in the exact rates of decline and in the asymptotes. This variation probably stems from interindividual variation in overall strength. Of greater importance in the present context, however, was the consistency of each individual's performance. Rarely did any subject show a greater spread in repeated contractions than is illustrated in Figure 3. This finding argues in favor of the use of constant-effort contractions as a dependent variable (Experiment I) and as an independent variable (Experiment II).

The effect of fatigue was apparent in two ways: (*a*) in the inability of some subjects to exert test squeezes at one or both of the two highest force levels, and (*b*) in the inflation of the estimates of intensity assigned to the test squeezes. Even when a subject seemed able to exert all of the force levels, his contractions at the two highest levels were occasionally of questionable validity. In their attempts to reach the highest levels, some subjects tended to jerk the handle of the dynamometer. Jerky motion, if applied with sufficient force, could produce a transverse movement of the arms of the handle and thereby bring the opposing capacitor plates into contact. The short circuit produced by contact of the plates might be only momentary, but even so, it called into question the force level actually achieved. Unfortunately, it was not always possible to correct this behavior during a session. Rather than eliminate only the questionable contractions from the data, it seemed prudent to eliminate the two highest force levels completely from further consideration.

The four sustained contractions in Group I all had about the same effect on the perceived intensity of the test contractions (Figure 4). In general, they caused a moderate, and rather uniform, ratio increase in perceived intensity. Power functions appear to fit equally

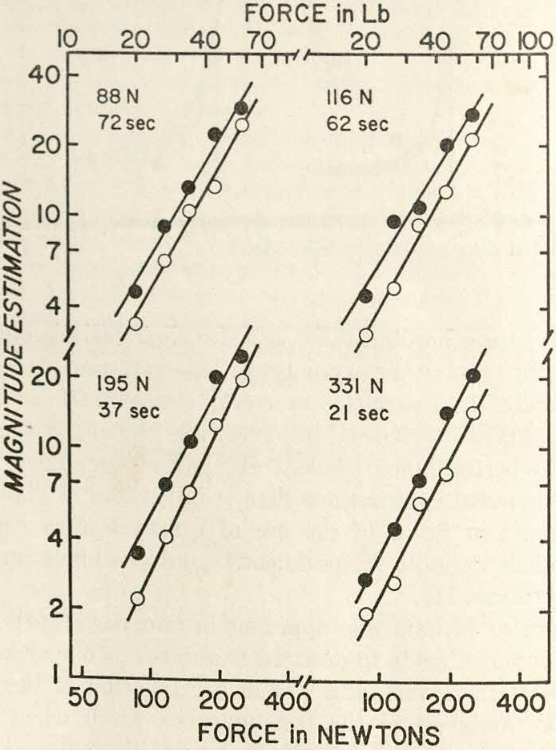


Fig. 4. Perceived intensity as a function of force for phasic contractions exerted after a nonfatiguing reference contraction (unfilled circles) and after a constant-effort, fatiguing, contraction (filled circles) at one of the four levels of initial intensity of Group I

well data obtained with and without prior fatigue. In addition, fatigue seemed to have little if any influence on the size of the exponent (slope). The *average* exponent obtained with fatigue was some 5% smaller than the average exponent obtained without fatigue.

Figure 5 shows the effects of the two contractions in Group II. Again, the estimates of apparent intensity were inflated, though surprisingly little more than the estimates after the Group I contractions. Power functions were again adequate to describe the data obtained with and without fatigue. The fatiguing stimulus whose initial intensity was 331 N. seemed to have negligible effect on the exponent, but the fatiguing stimulus whose initial intensity was 116 N. produced an exponent of 1.44, compared with 1.84 for the corresponding function without fatigue.

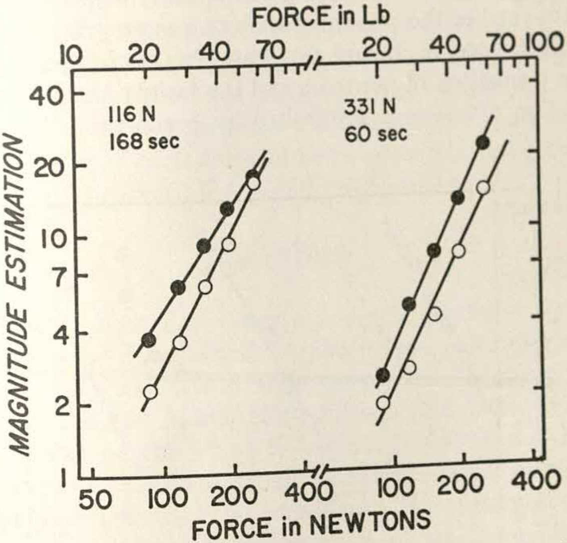


Fig. 5. Perceived intensity as a function of force for phasic contractions exerted after a nonfatiguing reference contraction (unfilled circles) and after a constant-effort, fatiguing, contraction (filled circles) at one of the two levels of initial intensity of Group II

The finding that all six of the fatiguing contractions produced very similar increases in the perceived intensity of phasic contractions casts doubt on the possibility that force level integrated over time might be used as an indicant of fatigue. When all six pairs of functions are viewed together, the constancy of the effect of fatigue on the exponent of the psychophysical function seems also more striking than the variation. (In the one case where fatigue seemed to reduce the exponent significantly, the fatiguing contraction was exerted over a long duration, 168 sec. Perhaps long contractions affect the exponent in a way not seen in contractions of moderate and short durations; it is also possible, however, that random variability in the intensity estimates accounts for the disparity.) In five of the six cases studied, fatigue did not significantly alter the exponent, and therefore its effect was virtually equivalent to multiplying the load of the test squeezes by an appropriate factor. This point is illustrated in Figure 6, where the nonfatigue results are plotted as a function of newtons, and the fatigue results are plotted as a function of newtons multiplied by a constant *W*. In log-log

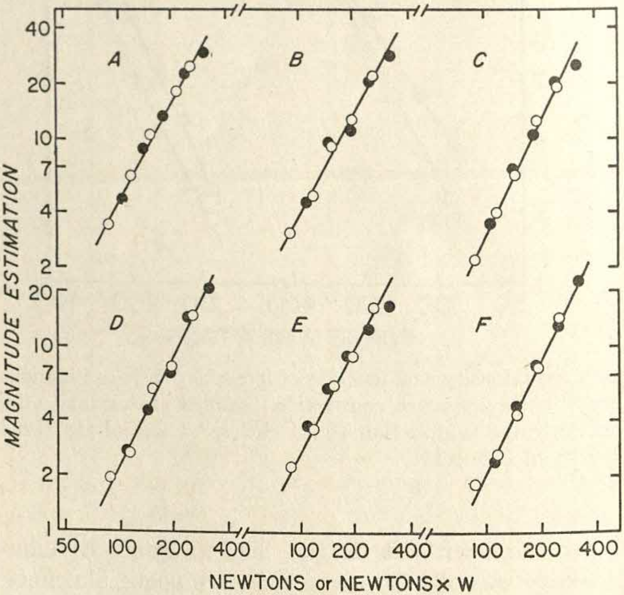


Fig. 6. Perceived intensity plotted as a function of newtons for the six nonfatigue conditions (unfilled circles) and as a function of newtons × *W* for the six fatigue conditions (filled circles). The straight lines are power functions fitted by the method of least squares to the unfilled circles. Curves A through D show the effects of the fatiguing contractions from Group I (88 N., 72 sec; 116 N., 62 sec; 195 N., 37 sec; and 331 N., 21 sec respectively) and curves E and F show the effects of the fatiguing contractions from Group II (116 N., 168 sec; and 331 N., 60 sec respectively).

coordinates, the size of W reflects the average horizontal difference between the power functions obtained with and without fatigue. The straight lines in Figure 6 are power functions fitted by least-squares to the data obtained without fatigue. The simple multiplicative transformation worked very well for five of the six functions, but less well for curve E (116 N., 168 sec).

The sizes of W for the Group I contractions were very similar: 1.20, 1.24, 1.24, and 1.23 for functions A, B, C, and D respectively. Exposure to each fatigue condition was therefore equivalent to a 20%–25% increase in the physical load of the test squeezes. The values of W for the Group II contractions were almost the same as for Group I contractions (1.22 for function E, 1.29 for function F). Of course, the use of the constant is not very meaningful for E. A W of 1.22 obviously 'overcorrects' the estimates at the highest two force levels, but a W of 1.30 would be necessary to bring the lowest three levels into line with the nonfatigue function.

DISCUSSION

The results of both Experiments I and II reinforce earlier evidence that the perceived effort of phasic and sustained handgrip contractions is amenable to straightforward mathematical description (see Stevens and Mack, 1959; Stevens and Cain, 1970). The need for two exponential terms to describe the constant-effort contractions suggests that two distinguishable processes may underlie the perception of effort. It is conceivable, for example, that one exponential term represents the activity of mechanoreceptors in the tendons or the skin, and the other term the activity of receptors sensitive to noxious metabolites produced by muscular activity. Both kinds of receptors can operate in sustained contractions, but it remains to be seen whether their roles can be separated at the psychophysical level. Some introspective evidence bears on this point: at the beginning of a contraction the sensation is localized mainly in the palm, and can be described as cramp-like, but as the contraction is continued (and the appropriate reduction in force has been made), the sensation tends to spread to the wrist and forearm, where the quality is experienced more as a 'burning.' Therefore, two qualities of sensation are felt by the subject and both are part of the total perceived effort.

It is, of course, arguable that results of curve fitting might bear little relationship to the number of processes or stages of a complex process that underlie the sensation. Often, alternative equations fit a given set of experimental observations. In the present case, however, the double-exponential function was found to describe seven sets of observations, and the values of the rate parameters remained invariant from one set of observations to another. This would seem to support the present description and to provide incentive to search further for some kind of dual mechanism in the perception of effort.

The invariance of the rate constants can be added to a growing list of invariances shown to characterize various muscle sensations (see Stevens and Cain, 1970). The exponent of the psychophysical power function for apparent effort appears to be nearly the same from one muscle group to another. The rate of change of the isometric endurance function (the function that relates the maximum length of time that a contraction can be exerted to the force level of that contraction) seems also to be invariant across muscle groups. For handgrip, the growth of apparent effort with duration does not depend on the force of contraction, and the growth of effort with force for contractions of various constant durations does not depend on the duration of the contractions. The evidence from Experiment II suggests that the exponent of the psychophysical power function for phasic contractions may also remain invariant from the nonfatigued to the fatigued state.

The simplicity and reliability of the constant-sensation procedure recommend it for the study of time-dependent sensory processes such as adaptation and fatigue. There has been little use of the method, although its advantages are several. One of its best points is that fatigue and adaptation can be charted at *any* level. Another is that the method seems well suited to the examination of the individual subject. In a study of adaptation in the thermal senses, it was found that a subject would reliably give the same adaptation curve when the task was to vary the temperature of a Peltier stimulator to maintain a minimal, but constant, sensation over time (Kenshalo and Scott, 1966). Consistent individual differences in the curves were also demonstrated.

Because of the suitability of the constant-sensation method for the study of the individual, the method may prove particularly useful in the examination of muscle disturbances and pathology. Taking a cue from Humphreys and Lind (1963), who studied how the endurance of a normal muscle compares with the endurance of a muscle whose blood supply is occluded, we have conducted pilot studies on how the constant-sensation function is affected by occlusion of the blood supply to the arm. Under occlusion, the early portion of the function is virtually the same as when circulation is normal. Later on, however, a dramatic difference appears. The occluded arm tires far more rapidly in the later stages, perhaps because of the accumulating noxious metabolites in the muscles. The important point here is how quickly and easily the constant-sensation procedure is able to detect a disturbance of muscle function in an individual subject.

Notes

This research was supported by Contract F44620-67-C-0017 with the United States Air Force Office of Scientific Research and was performed while the first author was a United States Public Health Service postdoctoral fellow. Received for publication April 23, 1970.

References

Bartoshuk, L. M. 1968. Water taste in man. *Perception and Psychophysics* 3:69–72.

Cain, W. S., and Engen, T. 1969. Olfactory adaptation and the scaling of odor intensity. In *Olfaction and taste III*, ed. C. Pfaffman, pp. 127–141. New York: Rockefeller University Press.

Eason, R. G. 1959. The surface electromyogram (EMG) gauges subjective effort. *Perceptual and Motor Skills* 9:359–361.

Humphreys, P. W., and Lind, A. R. 1963. The blood flow through active and inactive muscles of the forearm during sustained hand-grip contractions. *Journal of Physiology* 166:120–135.

Kenshalo, D. R., and Scott, H. A. 1966. Temporal course of thermal adaptation. *Science* 151:1095–1096.

McBurney, D. H. 1966. Magnitude estimation of the taste of sodium chloride after adaptation to sodium chloride. *Journal of Experimental Psychology* 72:869–873.

Stevens, J. C., and Cain, W. S. 1970. Effort in isometric muscular contractions related to force level and duration. *Perception and Psychophysics* 8:240–244.

Stevens, J. C., and Mack, J. D. 1959. Scales of apparent force. *Journal of Experimental Psychology* 58:405–413.

Stevens, J. C., and Stevens, S. S. 1963. Brightness function: Effects of adaptation. *Journal of the Optical Society of America* 53:375–385.

Stevens, S. S. 1960. The psychophysics of sensory function. *American Scientist* 48:226–252.

ADAPTATION TO UNIOCULAR IMAGE MAGNIFICATION AFTER VARYING PREADAPTATION ACTIVITIES

William Epstein
University of Wisconsin, Madison

Adaptation was measured after a 30-min exposure to a 5% meridional size lens, axis 90 deg. In the preadaptation period immediately preceding exposure, the subjects spent 10 min sitting in the dark with both eyes covered (C), monocularly exploring the building (M), or binocularly exploring it (B). The major finding was that the magnitude of adaptation was affected significantly by the preadaptation activity: C = M > B. This finding is discussed in the context of a general formulation of the process of adaptation to the lens.

A meridional size lens (MSL) worn before one eye magnifies the image in that eye. When the axis of the lens is vertical, the image is magnified along the horizontal meridian and the disparity between the images of the two eyes is altered, causing highly predictable perceptual distortion (Ames, 1946; Ogle, 1964). Epstein and Morgan have presented evidence (1970) that durative exposure to the lens results in adaptation to the disparity the lens produces. After binocularly exploring a typical indoor environment while wearing the lens, the subjects made depth judgments that exhibited significant shifts in an adaptive direction. The adaptive shifts were observed when the depth judgments were made in the absence of the lens (aftereffect), as well as when the lens was still in place (direct measure).

The adaptive shifts were attributed to a reevaluation of binocular disparity—to a recalibration of the relationship between disparity and perceived depth. The following was proposed as an account of the reevaluation process (Epstein and Morgan, 1970). When binocular disparity is paired with a competing set of monocular cues, a condition of informational discrepancy exists. The discrepancy is resolved initially by suppression of the information provided by disparity. In a visual field that contains the normal complement of monocular cues, suppression commences almost at once. The rapid onset of suppression is evident in the fact that even in the initial moments of exposure to the meridional size lens in a typical environment, the subject does not report full MSL effects. The suppression process sets the stage for the reevaluation of disparity. The reevaluation is assumed to be an associative process, proceeding along lines similar to the development of stimulus equivalence. That is, by pairing disparity with dominant monocular cues, disparity becomes equivalent to the monocular cues as a stimulus for depth.

The present experiment examined this hypothetical account by examining the effect of various preadaptation activities. The basic premise was that preadaptation activities that accelerate suppression of disparity also accelerate adaptation. Three preadaptation activities were compared: both eyes covered (C); undistorted monocular viewing (M); and undistorted binocular viewing (B). The sequence of events in the experiment was: first, an initial set of judgments of equidistance, with and without the meridional size lens; then, a 10-min interval filled by one of the three preadaptation activities; then, a 30-min period of exposure to the lens in a typical indoor environment; and last, a final set of equidistance judgments, with the lens. The difference between the initial and final equidistance judgments provided the measure of the adaptive shift that occurred during the 30-min period of exposure to the lens.[1] The principal objective was to determine whether the magnitude of adaptive shift would depend on the nature of the preadaptation activities that filled the second stage.

Undistorted monocular viewing (Activity M) introduces a period of disuse of binocular disparity, accompanied by forced reliance on monocular depth cues. This activity should reduce the length of time that is required for the suppression of disparity during the ensuing adaptation period. Consequently, given a relatively short adaptation period, the adaptive shift should be greater after Activity M than after Activity B. Undistorted binocular viewing (Activity B) provides exposure to disparity, as well as continued reinforcement of the preexperimental association between binocular disparity and the monocular cues; so suppression of disparity and the reevaluation leading to adaptation should be slower after Activity B. And when the eyes are closed (Activity C) there is again a period in which disparity is not used; so adaptation after Activity C should, it was expected, exceed adaptation after Activity

B. In addition, comparison of Activity C and M allowed us to determine whether forced exercise of monocular discrimination enhances adaptation over and above the effect of disuse of disparity.

METHOD

-Subjects-The subjects were 66 undergraduates with uncorrected vision. They volunteered to serve in the experiment to earn extra course credit.

-Apparatus-Two pairs of reversible melters' *goggles* were used. One pair had a 5% meridional size lens, axis 90 deg, in one frame and plain glass in the other frame. The other pair had plain glass in both frames. The *depth-discrimination apparatus* was identical to that used by Epstein and Morgan (1970); it is described in detail by Ebenholtz and Walchli (1965, pp. 455–456). Briefly, two fine electroluminescent vertical lines, separated laterally by 65 mm, are presented in total darkness. By turning a knob, the experimenter causes the two lines to move slowly and continuously, one advancing and the other receding by exactly the same amount. The lines are visible to the subject through a rectangular aperture at the face of the apparatus. The dimensions of the aperture ensure that the visible portions of both lines subtend the same, invariant, visual angle at all positions of the lines in the apparatus. When the subject's head is stationary and the room is dark, the discrimination of depth intervals between the lines depends exclusively on binocular disparity. Informal tests restricting the observer to monocular vision supported this claim by showing that depth discrimination was not possible without binocular vision.

-Procedure-The subject was seated with his head in a restraining device that prevented head movement. His head was frontal-parallel with the viewing aperture, and his eye-level was at the height of the midpoint of the aperture. The distance from his eyes to the equidistance position of the two test lines was 6 ft.

In the first stage of the procedure, four sets of four equidistance judgments were secured, using the method of adjustment. All subjects made the four sets of judgments in the same order: first, wearing the plain glass goggles; then, with the meridional size lens before 'other' eye; then, with the plain glass goggles again; and finally, with the lens before the eye that was slated to wear it during the adaptation period. Rest periods of 30-sec duration intervened between successive sets. During these intervals the goggles were removed, room illumination was increased to a dim level, and the subject was allowed to look about freely. The depth-discrimination apparatus was not visible. The test sequence required three changes of goggles; the changes were made while the subject's eyes were closed and without any indication by the experimenter about the nature of the glasses.

After the last set of judgments, the MSL goggles were removed and the subject participated in one of the three preadaptation activities. During Activity C, the subject sat at ease in the laboratory room, wearing a pair of opaque goggles that occluded vision completely. Activity M was a walk about the corridors of the Psychology Building. During the entire period one eye was occluded. Activity B was the same walk with both eyes uncovered. In each case the duration of activity was 10 min.

Upon conclusion of the 10-min period the MSL goggles were put on and a 30-min adaptation period began. During the adaptation period all subjects walked the corridors and stairs of the Psychology Building.

At the conclusion of the 30-min adaptation period, the subject was returned to the laboratory again. There he provided a final set of two depth judgments, wearing the MSL goggles.

RESULTS

The objectives of this experiment required subjects with uncorrected vision and good stereopsis. In a random sample of undergraduates in psychology courses, many students were found not to meet these requirements, especially for stereopsis. Three aspects of the preadaptation data were used to select subjects for participation: it was required (*a*) that there be no overlap among the 'plain-glass judgments' and the two sets of 'MSL judgments,' (*b*) that the mean plain-glass judgment not deviate by more than 1 cm from true equidistance, and (*c*) that the MSL judgments deviate from objective equidistance by at least 6 cm in the predicted direction. The ideal geometric effect for our experimental arrangement was 9.1 cm.[2] There were 18 subjects who failed to satisfy these criteria. They were terminated after completion of the first stage of the experiment. The data reported below are based on the performances of 48 subjects, 16 assigned to each preadaptation activity. Of these, 8 subjects in each condition adapted to the meridional size lens in front of the left eye, 8 to the lens in front of the right eye.

Preadaptation plain-glass judgments

Table 1 shows the means of the preadaptation plain-glass judgments (Pre PG). Since the variance was restricted sharply by our selection procedure, it is not surprising that there were no significant differences among the means.

Preadaptation MSL judgments

The meridional size lens produces an apparent rotation of the frontal plane, so that when the two test lines are objectively equidistant, the line on the same side as the lens appears farther back. Consequently, to achieve subjective equidistance, the right line must

Table 1. Mean separation in depth (cm) when test lines were judged equidistant

Preadaptation activity	Exposure eye		Combined means
	Left	Right	
Monocular (M)			
Pre PG	.81	.79	.80
Pre MSL	7.71	8.83	8.27
Post MSL	6.40	6.46	6.43
MSL diff	1.31	2.37	1.84
Closed (C)			
Pre PG	.71	.59	.65
Pre MSL	8.11	7.78	7.94
Post MSL	6.28	5.68	5.98
MSL diff	1.83	2.10	1.96
Binocular (B)			
Pre PG	.67	.96	.81
Pre MSL	7.01	6.78	6.90
Post MSL	6.47	5.65	6.06
MSL diff	.54	1.13	.84
Combined MSL diff	1.23	1.87	

be moved forward when the right eye is exposed to the lens, and the left line must be moved forward when the left eye is exposed to the lens. As noted earlier, the predicted separation was 9.1 cm. Table 1 (Pre MSL) shows the obtained mean separation in depth when the two test lines were judged equidistant. The means for the six groups ranged from 6.78 to 8.83 cm.

An analysis of variance was computed to determine whether the means associated with the three preadaptation activities differed and whether there was a difference associated with exposure eye. No significant effects were revealed. Therefore, we may conclude that the subjects assigned to the three activities were equally susceptible to the meridional size lens. In addition, we may conclude that both eyes yielded comparable MSL effects prior to the adaptation period.

Postadaptation MSL judgments

Table 1 (Post MSL) shows the mean postadaptation MSL settings. Adaptation to the lens is evident when the mean separation in depth is smaller following exposure than preceding exposure. This was the case for all three conditions of preadaptation activity and for both eyes. For each subject, the difference between the preadaptation and postadaptation depth interval was computed. A positive sign was assigned to the difference when the postadaptation interval was smaller, and a negative sign when the postadaptation interval was larger. Of the 48 subjects, 40 showed positive shifts, 7 showed negative shifts, and 1 did not shift.

A constant was added to each difference score, which eliminated the negative signs, and an analysis of variance was computed on the transformed differences. The effects of preadaptation activity $[F(2, 42) = 4.06]$ and exposure eye $[F(1, 42) = 4.76]$ were significant ($p < .05$). The interaction was not significant. Inspection of Table 1 reveals the nature of the significant effects. Activities C and M yielded comparable pre–post differences (i.e., adaptive shifts), but both C and M yielded greater differences than Activity B. For each activity, the difference was greater when the right eye was exposed to the meridional size lens.

Table 2. Mean percent adaptation after different preadaptation activities

Preadaptation activity	Exposure eye Left	 Right	Combined means
Monocular (M)	18.0	30.7	24.3
Closed (C)	21.7	29.1	25.4
Binocular (B)	8.5	15.0	11.7
Combined means	16.1	24.9	

Percent adaptation

The percent adaptation was computed for each condition using the formula

$$\frac{\text{Pre MSL} - \text{Post MSL}}{\text{Pre MSL} - \text{Pre PG}} \times 100$$

These data are presented in Table 2. The significant effects revealed by the preceding analysis are discerned readily in Table 2. Percent adaptation was least for Activity B; Activities M and C differed very little. For each activity, percent adaptation was greater when the right eye was exposed to the lens. Analysis of variance of the percent-adaptation data showed that the main effects, of preadaptation activity [$F(2, 42) = 3.98, p < .05$] and exposure eye [$F(1, 42) = 7.32, p < .01$], were both significant.

DISCUSSION

The experiment was designed to examine an implication of Epstein and Morgan's account (1970) of the process responsible for adaptation to a meridional size lens. The account implied that activities immediately preceding the adaptation period would affect the speed of adaptation. Activities that encourage reliance on monocular depth cues were expected to accelerate suppression of disparity and to facilitate adaptation. Monocular exploration of a setting that contains the typical complement of depth cues is an obvious example of this type of activity. On the other hand, activities that reinforce the preexperimental relationship between disparity and depth, and between disparity and monocular depth cues, were expected to retard adaptation. Unrestricted binocular exploration of the typical extralaboratory environment is an example of this type of activity. The experimental data confirmed these expectations. The 30-min adaptation period yielded 24.3% of full adaptation when it was preceded by Activity M, and only 11.7% of full adaptation when it was preceded by Activity B.

Activity M was a period of forced reliance on monocular depth cues, but it was also a period of disuse of binocular disparity. Although our discussion of Activity M has stressed the former factor, it is obvious that the facilitative effect of Activity M could be due to disuse of disparity. The results for Activity C help resolve this ambiguity. Adaptation following Activity C exceeded the magnitude of adaptation that followed Activity B and did not differ significantly from the effect obtained after Activity M. Similar results for preadaptation with the eyes closed, as in Activity C, have been found in recently completed studies of adaptation to optical tilt.[3] The fact that Activities C and M had comparable effects suggests that the effective distinction between these activities and Activity B is their disuse of binocular disparity. Apparently, a period of disuse weakens the relationship between disparity and perceived depth.[4] Weakening of this relationship is conducive to suppression of binocular disparity and more rapid adaptation. The facilitative effect of disuse is not enhanced by the forced exercise of monocular discrimination, as in Activity M.

The experimental data also show significant effects associated with exposure eye. Adaptation was greater when the right eye was exposed to the meridional size lens. This difference cannot be attributed to differences in initial susceptibility to the lens, since the preadaptation effects did not differ for right- and left-eye exposure to the lens. Nor was the difference due to an artifact of our experimental procedure. A similar difference in favor of the right eye has been reported by Burian (1943), using a variety of procedures, and we have found the difference with a procedure adopted from Gillam (1968). At present we have no explanation to offer for this difference. Burian suggested that the difference may be related to ocular dominance, but he presented no evidence for this proposal. We made informal assessments of ocular dominance with the help of a pointing task and failed to uncover the hypothesized relationship.

Notes

This investigation was supported by Research Grant MH-1630-02 from the National Institute of Mental Health. Judith Turner provided valuable assistance in all phases of this investigation. Received for publication June 4, 1970.

1. Most contemporary investigations of adaptation have relied exclusively on the aftereffect measure, instead of the direct measure used in the current study. The relative merits of the two measures have been discussed by Epstein (1967, Ch. 9) and by Rock (1966, Ch. 1). More recently a number of investigators have reported puzzling discrepancies between the two measures and have questioned the equivalence of the direct and aftereffect measures (see Freedman, Wilson, and Rekosh, 1968; Foley and Abel, 1967; Foley and Maynes, 1969; and Canon, 1970). Most relevant for consideration here is the fact that although both the direct and the aftereffect measures showed significant adaptive shifts in Epstein and Morgan's study of adaptation to the lens (1970), the magnitudes of the adaptive shifts yielded by the two measures were not correlated significantly. For these reasons, as well as considerations of experimental convenience, the aftereffect measure was omitted from the present study.
2. For a meridional size lens at axis 90 deg, the equation for predicting the apparent rotation of the frontal plane is $\tan \psi = [(M - 1)/2M]\ [y/a]$, where M is lens magnification as a percentage, y is observation distance, and a is one-half the interocular distance. For a discussion of this equation, see Ogle (1964, pp. 160–166).
3. S. M. Ebenholtz, manuscript in preparation.
4. Evidence that disuse of disparity leads to fast deterioration of the effectiveness of disparity is presented by Wallach and Karsh (1963).

References

Ames, A., Jr. 1946. Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments. *American Journal of Psychology* 59:333–357.

Burian, M. 1943. Influence of prolonged wearing of meridional size lenses on spatial localization. *Archives of Ophthalmology* 30:645–666.

Canon, L. K. 1970. Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *Journal of Experimental Psychology* 84:141–147.

Ebenholtz, S. M., and Walchli, R. M. 1965. Stereoscopic thresholds as a function of head and object orientation. *Vision Research* 5:455–461.

Epstein, W. 1967. *Varieties of perceptual learning.* New York: McGraw-Hill.

Epstein, W., and Morgan, C. L. 1970. Adaptation to uniocular image magnification: Modification of the disparity-depth relationship. *American Journal of Psychology* 83:322–329.

Foley, J. E., and Abel, S. M. 1967. A study of alternation of normal and distorted vision. *Canadian Journal of Psychology* 21:220–230.

Foley, J. E., and Maynes, F. J. 1969. Comparison of training methods in the production of prism adaptation. *Journal of Experimental Psychology* 81:151–155.

Freedman, S. J., Wilson, L., and Rekosh, J. H. 1968. Three different methods for producing shifts in visually guided reaching after exposure to optical displacement. In *Neuropsychology of spatially oriented behavior,* ed. S. J. Freedman. Homewood, Ill.: Dorsey Press.

Gillam, B. 1968. Perception of slant when perspective and stereopsis conflict. *Journal of Experimental Psychology* 78:299–305.

Ogle, K. N. 1964. *Researches in binocular vision.* New York: Hafner.

Rock, I. 1966. *The nature of perceptual adaptation.* New York: Basic Books.

Wallach, H., and Karsh, E. B. 1963. Why the modification of stereoscopic depth-perception is so rapid. *American Journal of Psychology* 76:413–420.

DISCONTINUITIES IN RECOGNITION LEARNING REVEALED BY CRITICAL-TRIAL ANALYSIS

Robert Teghtsoonian and Martha Teghtsoonian
Smith College

The number of choices required to locate a previously seen item among a group of alternatives (three-digit numbers or consonant trigrams) was found to be a discontinuous function of successive viewings. Until the critical trial—the first occasion on which the correct item is the subject's first choice—performance was only slightly better than chance. Immediately after the critical trial, the mean number of choices moved to and remained at a sharply lower level. The discrete transition at the critical trial was concealed by conventional averaging over trials numbered from the beginning of the task but emerged clearly when averaging was performed over trials realigned with the critical trial as reference point for each item.

When subjects are asked to locate a previously seen item by choosing from a set of alternatives until successful, the number of choices required changes discontinuously over trials. R. Teghtsoonian described such a task in which each choice was made from a group of 16 three-digit numbers, and in which each trial included 12 such groups each containing just one previously seen item (1964). Until the *critical trial*—the first occasion on which a particular item was identified with the first choice—the number of choices required to locate the item was fairly stable at about seven. Immediately following the critical trial, the number of choices dropped nearly to one and remained low on succeeding trials. This strikingly discontinuous pattern was observed for all subjects, and indeed was apparent in the data for single items.

Since more critical trials occur early in practice than late, the average number of choices at each trial reveals the gradual decline typical of continuous acquisition functions. For group averages to display the discrete pattern characteristic of individual items, it is necessary to realign trial sequences so that the critical trial for each item becomes the reference point, and to calculate averages for trials numbered relative to that reference point. Thus, two alternative acquisition functions can be generated from the same set of data. The customarily averaged performance measure, plotted against trials numbered from the beginning of the task, shows a regular and gradual decline in the number of errors. The other function, obtained by averaging after trials have been renumbered relative to the critical trial, shows clearly at that point a sharp discontinuity in error frequency.

The present investigation was planned to check the generality of the original finding of discontinuity in the error function around the critical trial with a larger sample of subjects that had been drawn from a different population, and to extend the repeated-choice technique to a broader sampling of experimental materials. Accordingly, one group of subjects was assigned the same problem as in the earlier experiment: learning to recognize a group of 12 three-digit numbers. For two additional groups, the three-digit numbers were replaced by low-frequency and by high-frequency consonant trigrams respectively.

Further, the previous experiment had shown that prior to the sharp reduction in errors at the critical trial, the subjects were not totally ignorant of the identities of the correct items; while error levels were high, they reflected better performance than would result from guessing. And following the critical trial, while errors remained rare, some items did again require more than one choice. The present experiment was planned with a view to a closer examination of both the precritical-trial learning and the postcritical-trial errors.

METHOD

–The task–Each subject was told that he would be shown a set of items, one of which had been designated 'correct,' and that he was to guess until informed that he had found it. Eleven additional sets of a similar nature, each bearing a single correct item and 15 different alternatives, completed Trial 1. The subject was informed that each subsequent trial would include the same 12 correct items but that their order and spatial location in each display, and the identities of the incorrect alternatives, would vary. He was instructed to proceed until he could complete a trial without error, that is, could identify each correct item in a single guess.[1]

-Materials-One group of items to be learned was composed of 12 randomly selected *three-digit numbers*. Trial 1 was constituted by grouping each of the correct items with 15 alternatives; in selecting the total of 180 incorrect alternatives for Trial 1, sampling was performed randomly, and without replacement in each set of 15. Four additional trials were prepared in the same manner, preserving the identities of the 12 correct items as the only factor common to all trials. Over the five trials, the order of the correct items, their location within a set of alternatives, and the identities of the incorrect alternatives were all determined randomly. If more than five trials were required to achieve an errorless trial, the sequence was repeated from Trial 1. The task was self-paced; a new card was presented only after the subject identified a correct item. All responses were given verbally and were recorded by the experimenter; each response was immediately identified for the subject as correct or incorrect.

Another group of items to be learned was composed of *consonant trigrams*. The 10 *least frequently occurring* consonants in printed English text were identified as B, G, J, K, P, Q, V, W, X, and Z, and in that alphabetical ordering were mapped into the digits ordered from zero through nine. According to this code, the entire set of materials described above was transformed into a new set, formally identical with the first but substituting the least frequent consonants for the digits. The *most frequently occurring* consonants were identified as C, D, F, H, L, M, N, R, S, and T, and following the same coding procedure, a third set of materials was produced.

-Subjects-There were 16 subjects in each of the experimental conditions. All 48 were summer students at the University of Massachusetts and were paid for their services.

RESULTS AND DISCUSSION

Conventionally averaged error functions

In order to provide a standard of comparison for subsequent analyses, the mean number of incorrect choices, or errors, was determined for each subject for every trial numbered from the beginning of practice. The means of these values for the three conditions are shown in Figure 1. All three error functions exhibit the gradual reduction typical of this method of averaging. Incidentally, they indicate that the three-digit numbers and the high-frequency consonant trigrams were very similar in difficulty, while the low-frequency consonant trigrams appear to have been appreciably harder to master. A one-way analysis of variance for total errors yielded $F = 2.59$, $df = 2, 45$, $.05 < p < .10$. This is consistent with the finding that while the numbers of subjects who failed to reach criterion for the three-digit numbers and the high-frequency consonant trigrams were 1 and 2 respectively, 6 of the 16 subjects presented with low-frequency consonant trigrams failed to reach criterion.[2]

Error functions averaged over trials numbered relative to the critical trial

The first trial on which an item was correctly located with the first choice was called a critical trial, and all other trials were renumbered relative to it. Thus, the three trials immediately preceding the critical trial were renumbered *−3*, *−2*, and *−1*, while the three succeeding trials were renumbered *+1*, *+2*, and *+3*. All averages were calculated for these renumbered trial values, whether over items for a particular subject, or over subjects for a particular group.

For each subject the mean number of errors was calculated for the three trials preceding and following the critical trial,[3] and the means of these values for each group are shown in Figure 2 as filled circles. In all three conditions there is the same abrupt drop in mean error level following the critical trial that was observed in the previous study (Teghtsoonian, 1964). It is important to note that the nonzero values which follow the critical trial are a composite of the many cases in which the error level remained at zero and those cases in which errors did recur. For 29 of the 48 subjects, at least half of the items involved no errors on the first trial following the critical trial; a majority of the items fall into this category and of course their mean error level is zero.[4]

For the remaining items, a separate calculation was made of the mean error level, and these values, shown as unfilled circles in Figure 2, reveal an interesting pattern: when items were 'forgotten' following the critical trial, the number of choices required rose to a level apparently continuous with the number of choices required prior to the critical trial. It seems that items are either retained in an errorlessly available state or revert to a state in which several choices are required for their location, the number of choices beginning at a level close to that prevailing prior to the critical trial and declining in a gradual fashion. The data of the previous study (Teghtsoonian, 1964) were reexamined for evidence of this particular phenomenon, and the results, shown in Figure 3, indicate a similar pattern.

The similarities between Figures 2 and 3 may be summarized as follows. First, all groups exhibit a sharp discontinuity at the critical trial. In a conventional data analysis, this drop in the number of

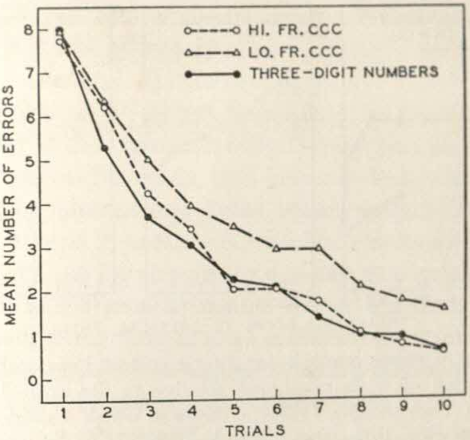


Fig. 1. Number of errors made in locating a correct item among 16 alternatives, as a function of trials numbered conventionally from the beginning of the task

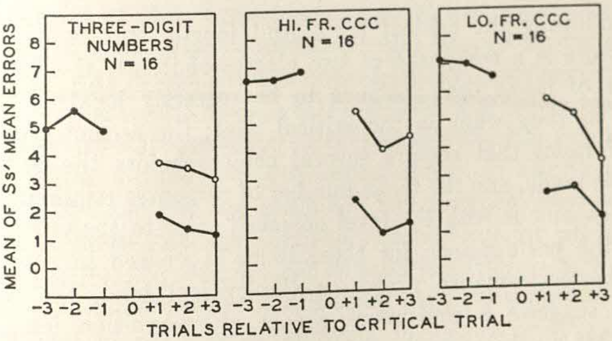


Fig. 2. Number of errors made in locating a correct item among 16 alternatives, as a function of trials numbered relative to the critical trial for each item. The filled circles represent all items. The unfilled circles represent items which, after the critical trial, required more than a single choice to locate.

errors is concealed because the location of critical trials varies over items and subjects. Second, in all cases the level of performance prior to the critical trial is better than chance.[5] Thus, before the sharp improvement that occurs at the critical trial, the subjects have, on the average, learned something that enables them to locate the correct items with fewer choices than would be required by guessing. Third, in all cases the number of choices required remains at a low

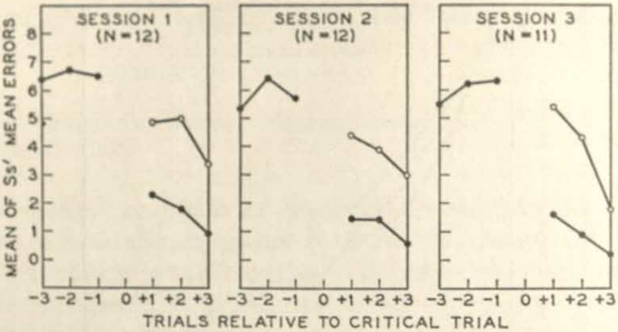


Fig. 3. Number of errors made in locating a correct item among 16 alternatives, as a function of trials numbered relative to the critical trial for each item. Data from Teghtsoonian (1964), a study in which a single group of subjects ($N = 12$) served for three sessions; the materials were three-digit numbers. The filled circles represent all items. The unfilled circles represent items which, after the critical trial, required more than a single choice to locate.

level following the critical trial. And fourth, this relatively low mean score is a composite of two classes of items: the first group consists of items that continue to be correctly located with one choice, as they were on the critical trial; the second group consists of items that require several choices before the correct response is made, and the mean number of responses required appears to be continuous with the level achieved prior to the critical trial. When the performance for these items is viewed in conjunction with the performance preceding the critical trial, the resulting picture suggests a continuous aspect of recognition learning. It seems that a subject gradually improves in his ability to reduce the number of incorrect choices and that during the process, items frequently make a sudden transition from a state in which several choices are required to one in which no errors are made.

Learned reductions in error-run lengths

Perhaps more puzzling than the discrete change in performance levels at the critical trial is the evidence of learning prior to the critical trial and the still-better performance levels for items involving the recurrence of error following the critical trial. It is not clear just what has been learned. One possibility is that the subject is learning to identify incorrect alternatives as such and that this skill may be exercised for a set of alternatives even when the correct item is not immediately recognized. The subject may have learned some general properties of the list of correct items (e.g., no multiples of 100, no item beginning with a particular digit) even though unable to correctly recognize each item on that list; on that basis he is able to exclude certain alternatives from consideration. This 'exclusion learning' hypothesis is consistent with the results of an analysis of error popularity for low-frequency consonant trigrams: 8% of the errors on Trials 2–6 were accounted for by the three least popular choices on each card, compared to a chance expectation of 20% of the errors. The subject does appear to avoid certain incorrect alternatives even when he is mistaken in his choice.

The same error-popularity analysis of the low-frequency consonant trigrams also suggests that some of the incorrect alternatives are chosen more frequently than expected by chance. On Trials 2–6, 33% of the errors were accounted for by the three most popular choices on each card. It seems likely that the subject develops partial knowledge of some correct items and on that basis selects particular alternatives. This conclusion is also implied by the relative frequencies with which error sequences of various lengths occur. For all three tasks, the modal error sequence is of length one. That is, the single most frequently occurring error is the case in which the correct item is located with the second choice. If error sequences of lengths one and two are combined and summed over the three tasks, they constitute the following proportions of all subject-item errors: .24, .27, .37, .42, .48, and .52 for Trials *−3*, *−2*, *−1*, *+1*, *+2*, and *+3* respectively. (In every case, guessing would yield a proportion of .13.) Even allowing for short error sequences occurring by chance, the subject appears to be choosing in many cases from among only two or three alternatives. It seems unlikely that such a reduction of the number of effective alternatives is achieved by exclusion alone.

Although both exclusion learning and partial learning of correct items are possible explanations for the better-than-chance performance levels when errors occur, we are not able to draw any conclusions about their rate of growth. Although the error functions in Figure 2 suggest gradual improvement, this may be due to the same sort of artifact that generates the gradual decline of the error functions in Figure 1. The present experiment was designed around the concept of learning to recognize a complete item, and we could readily stipulate the behavioral criteria for a critical trial. Different procedures would be required for the definition of critical trials in the cases of exclusion learning and partial learning. It remains to be seen whether such critical trials would mark similar discontinuities to those observed for whole-item recognition.

SUMMARY AND CONCLUSIONS

The present results confirm and extend the previous finding that the first errorless selection of a correct item from a group of alternatives marks a discrete transition in the subject's capacity to recognize previously seen items. Prior to that critical trial, the subject requires several choices before making the right one, but afterward, the number of choices is near zero for most items. This outcome has now been observed for two kinds of consonant trigrams as well as for three-digit numbers.

However, the sharp transition at the critical trial does *not* proceed from guessing to consistently errorless performance. Prior to the critical trial, the subject has learned enough about the identities of the correct items to reduce the lengths of his error sequences below the guessing level. And following the critical trial, some items do revert to a state requiring several choices with an average value that appears continuous with levels prevailing prior to the critical trial. Such a finding is consistent with a model proposing a process of gradual learning during which sudden transitions to an errorless state occur as critical thresholds are attained for particular items. But the gradual aspect of the error functions may reflect the same kind of statistical artifact that is associated with conventionally averaged error data. Whatever the subject is learning aside from the actual identities of the correct items, such learning may itself involve critical trials and may be characterized by discontinuities at that point. Of all the things that the subject may be learning, we are able to define critical trials only for correct item location, and when we do, the associated discontinuities are evident. Perhaps a similar outcome would obtain if other dimensions of the subject's learning, and the associated critical trials, could be as precisely defined.

Notes

This research was supported by USPHS Grant MH-11809-01. The authors thank S. Sternberg for valuable discussions of these data, and the Behavioral

Research Laboratory of the Bell Telephone Laboratories, Inc., at Murray Hill, New Jersey, where a first draft was written, for the use of their facilities. We also thank Anne Knecht Wilson for her help in running subjects and tabulating data. Received for publication July 2, 1970.

1. In fact, we report below data for the first 16 subjects run under each condition, even though this meant including 9 subjects who did not reach criterion. These subjects were unable to continue to the end of the procedure because of pressures of time; the shortest period of time served by any of these 9 was 80 min.

2. Despite these failures, the total number of subject items for which a critical trial could be determined was very similar in all three conditions: 190, 185, and 190, for three-digit numbers, low-frequency consonant trigrams, and high-frequency consonant trigrams respectively, out of a maximum possible of 192 (12 items for each of 16 subjects). Further, the number of trials run was similar among the three conditions. The mean values were 11.56, 12.31, and 10.56 for three-digit numbers, low-frequency consonant trigrams, and high-frequency consonant trigrams respectively; a one-way analysis of variance yielded $F < 1.0$ for $df = 2, 45$.

3. Trial 1 scores were never included in this analysis since they could not reflect any effects of learning. Occasionally, the critical trial occurred within the first four or last three trials. For such items, therefore, data were not available for all six trials relative to the critical trial. Such cases were, however, included in the calculation of averages, but for each subject so affected, his mean scores were calculated for the reduced number of items. While it was common for critical trials to occur early, very few occurred within the last three trials (due to the fact that the end of the task was commonly delayed by the recurrence of errors for items after a critical trial had been achieved). A separate analysis for the effect of critical-trial number suggests that the pattern of discontinuity at the critical trial occurs at all stages of the task.

4. This fact can be given greater prominence by representing each subject's performance by medians rather than means, as was done in the previously reported study. However, this has the disadvantage of obscuring the relatively high error levels characteristic of those items involving the recurrence of error.

5. Each subject was represented by the mean of his error frequencies on Trials *−1*, *−2*, and *−3*, and his score was subtracted from 7.5, the level expected by chance. The resulting difference distributions yielded values of $t = 7.72$, 1.74, and 1.61 for three-digit numbers, low-frequency consonant trigrams, and high-frequency consonant trigrams respectively (in each case, for $df = 15$). The mean differences were reliably greater than zero with $p < .001$ for three-digit numbers, but $.05 < p < .10$ for both forms of trigrams. A one-way analysis of variance performed on the three distributions showed a reliable difference among them ($F = 8.34$, $df = 2, 45$, $p < .01$); pair-wise comparisons showed that three-digit numbers differed reliably from both varieties of trigram but that the latter did not differ between themselves. The better-than-chance level observed previously for three-digit numbers is confirmed here, but the level was clearly superior to that obtained for either variety of trigram. Whatever the subject learns prior to the critical trial is

learned to a greater extent for three-digit numbers. This difference occurred despite the formal identity in the structures of items for all three types of material, and it poses some interesting questions about the variables which control the level of performance that precedes the critical trial.

References

Teghtsoonian, R. 1964. One-trial learning directly observed. *Canadian Journal of Psychology* 18:304–310.

EVENT TRACKING IN PROBABILITY LEARNING

Arthur S. Reber and Richard B. Millward
University of British Columbia and Brown University

After 1,000 observation trials in which the event schedule followed a sawtooth wave with a period of 50 trials, subjects tracked the events very closely on periodic test trials but showed no periodic tracking on no-event or constant-probability test trials. During the early periodic-event cycles the subjects' response proportions tracked the event proportions approximately as predicted by statistical learning theory, but during later cycles, their response proportions peaked and bottomed at the same points as the event proportions, a relationship nicely described by the averaging model of Friedman, Carterette, and Anderson.

Only two studies have been reported in which the sequence of reinforcing events in a probability-learning experiment was systematically varied according to some long repetitive function. Estes (1957a) mentioned an unpublished study in which the event probabilities varied according to the contingent linear function $a_1 + b_1n$, if an A_1 response was made, and $a_2 + b_2n$, if an A_2 response was made. In that study, he was testing a development of statistical learning theory for the changes in behavior with constant, variable, and contingent probabilities of reinforcements (1957b). The second study (Young, 1967) indicated that when the event sequence is sinusoidal, the subjects' responses reflect the periodicity. Both studies give qualitative, but not rigorous, support to statistical learning theory. However, both contained far too few trials to measure asymptotic behavior reliably. The primary theoretical purpose of this experiment was to test statistical learning theory for long-term, stable, asymptotic behavior with varying event sequences.

The present experiment was also run to provide evidence on several empirical issues critical for any theory of probability learning. The first of these concerns the extent to which subjects' response probabilities are sensitive to events occurring on previous trials. Millward and Reber (1968) and Vitz and Hazan (1969) have both shown that with Bernoulli event sequences subjects attend to and retrieve information from rather remote trials. It seemed important to study remote dependencies with structured event sequences.

A second empirical concern was with the extent to which subjects detect and respond to a well-defined cycle that extends across many trials. Anderson (1964), in his criticism of statistical learning theory, argued that long-term event patterns influence response probability. The probabilistic event cycle used in the present study seemed appropriate for testing this notion.

The third empirical concern was with the role of no-event trials. Previous research (Reber and Millward, 1968) has indicated that with Bernoulli sequences, response proportions continued to approximate π even after the events were no longer presented. It was considered important that we try to evaluate the statistical nature of responding on these 'blank' trials when the event sequence had long-term structure.

METHOD

-Procedure-Three types of trials were involved in the experiment: *observation, no-event*, and *prediction* trials. Observation trials consist of the presentation of events at a rate of two per sec. The subject does not respond during these trials and no preparatory signal is presented; he simply observes events according to some presentation schedule (see Reber and Millward, 1968, for a full discussion of this procedure). The no-event, or 'blank' trial, procedure has a long history in the literature on probability learning (e.g., Millward, 1967; Greeno and LaBerge, 1963). On no-event trials, a signal occurs indicating that the subject should respond, he responds, but no event occurs. The instructions imply that the events have been programmed in a manner statistically similar to those the subject has just seen but that he will not see them. He is told to respond on each trial just as if an event has occurred and to be correct if he can. Prediction trials are standard probability-learning trials with a signal to respond, a response, and a reinforcing event.

The experiment lasted three days. On each day, the subjects received 1,000 observation trials followed by 1,000 test trials. These test trials constituted an independent variable of the experiment. In all cases, the event sequence for the 1,000 observation trials was defined according to the functions

$$\pi_n = \begin{cases} \bar{\pi} + S(n-1) & \text{for} \quad 1 \leq n < P_1 \\ \bar{\pi} - S(n-P_2) & \text{for} \quad P_1 \leq n < P_3 \, , \\ \bar{\pi} + S(n-P) & \text{for} \quad P_3 \leq n \leq P \end{cases} \qquad [1]$$

where $P_1 = (P+3)/4$; $P_2 = (P+1)/2$; $P_3 = (3P+1)/4$; $S = 4\sigma/(P-1)$; π_n is the probability of an E_1 event on Trial n; and the index n is taken modulo P. The parameters $\bar{\pi}$, P, and σ define the mean E_1 event probability, the period of the saw-tooth wave, and the excursion from $\bar{\pi}$ respectively. In this experiment $P = 50$, $\sigma = .4$, and $\bar{\pi} = .60$. Thus, the wave started on Trial 1 at .6, increased linearly with constant $S = .03265$ until Trial 13 ($\pi_{13} = .9918$), decreased linearly until Trial 38 ($\pi_{38} = .2082$), and then increased again to .6 on Trial 50. This cycle was repeated for 20 blocks of 50 trials.[1] The actual averaged event proportions $[P(E_1)]$ realized by the random events generated with Equation 1 on Day 1 are shown in Figure 1b.

The presentation of the 1,000 observation trials took approximately 8 min. The purpose of these trials was to ensure that the subjects were at asymptote prior to running the test trials (see Reber and Millward, 1968). On Day 1, following the observation trials, the subjects received 1,000 standard prediction trials on which the event sequence was the wave function defined in Equation 1. On Day 2, the test trials were 1,000 no-event trials, while on Day 3, the test trials were 1,000 prediction trials with a constant probability .60 of E_1 event. There were no special instructions on the nature of the event sequences.

-Subjects–The subjects were 20 high-school juniors and seniors from the Providence school system. They were run for two hours a day for three days and paid $7.50. The 20 subjects were run in five subsets of four subjects each. All subjects in a subset had the same sequence of observation and event outcomes. The timing of events and signals was as described in an earlier paper (Reber and Millward, 1968).

-Apparatus–The subjects were seated in four separate booths. In front of each subject was a response panel with two push buttons and two event lights. Centered above the event lights was a signal light. The experimenter sat behind the subject and controlled the presentation of the events.

RESULTS

Day 1

On Day 1, the reinforcing events on test trials followed the same function that defined the events of the observation trials. Subjects systematically tracked the event sequence as it varied over each repetition of the wave pattern. Figure 1a shows a comparison between the theoretical event probabilities $[P(E_1)]$ and the proportions of A_1 responses $[P(A_1)]$, in blocks of five trials from both the initial two cycles and the final two cycles of the test trials. Note there that although the 1,000 observation trials had produced some learning about the nature of the event sequence, the subjects were not perfectly tracking the wave on Cycles 1–2. Initial tracking has generally smaller amplitude than asymptotic tracking, and the peaks and troughs on Cycles 1–2 show considerable lag behind the $P(E_1)$ curve as compared with Cycles 19–20.

The details of overall trial-by-trial tracking are shown in Figure 1b. Here $P(A_1)$ on the nth trial within each 50-trial block is plotted and compared with $P(E_1)$. Overall $P(A_1)$ was .627, or slightly over the $\bar{\pi}$ value of .60. The response proportions and the observed event proportions differ in variability and amplitude. The larger variance in the event proportions is probably due to the fact that only 50 realizations of the event sequence are averaged in each point, while 200 subjects' proportions are averaged for each point of the response curve. The maximum and minimum response proportions were .84 on Trials 13, 14, and 15, and .39 on Trial 37. According to the theoretical event schedule, the maximum and minimum event proportions occur on Trials 13 and 38, with values of .99 and .21 respectively. Thus, the overall response proportions showed effectively no lag with respect to the event proportions.

Another way to illustrate what the subjects were doing is to look at response proportions on odd-numbered and even-numbered 25-trial blocks. According to the definition of the event schedule, an overall $P(E_1)$ of .80 is expected in the first (odd) 25 trials and an overall $P(E_1)$ of .40 is expected in the second (even) 25 trials of each 50-trial block. Figure 2 shows that the event proportions were about as expected and illustrates the closely matching behavior observed in the response proportions. Across the response and event curves, almost every rise and dip in the event proportions is matched by a rise and dip in the response proportions. There was a closer match between $P(A_1)$ and $P(E_1)$ on the odd 25-trial blocks than on the even 25-trial blocks, as was expected from the observed overshooting of $\bar{\pi}$. The average response proportions for the odd and even trial blocks were .739 and .516 respectively. Subjects were obviously following the event sequence very closely.

Day 2

On Day 2, the subjects received 1,000 observation trials generated according to the wave schedule, and then 1,000 no-event trials on which they were asked to predict without event feedback. If one thinks of the observation trials as continuing into the pre-

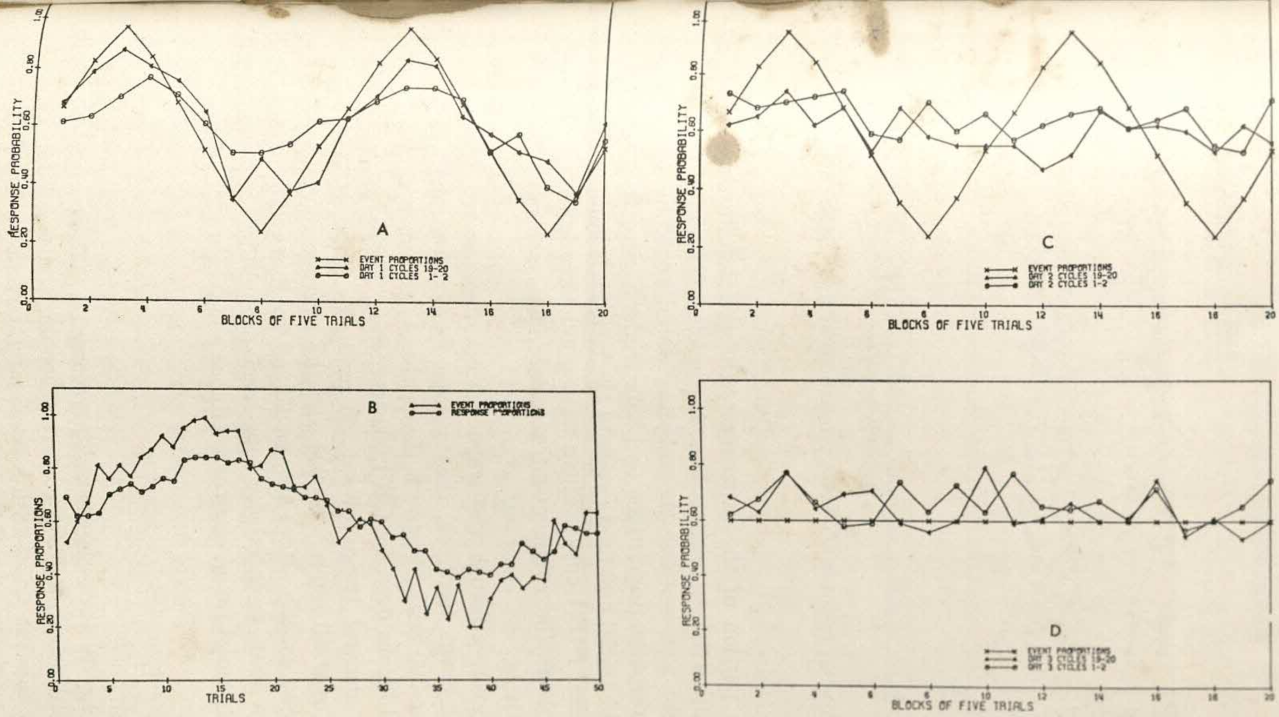


Fig. 1. (a) Theoretical event probabilities and response proportions in blocks of five trials for Cycles 1–2 and Cycles 19–20 on Day 1. (b) Average trial response and event proportions over all 20 blocks of 50 trials each on Day 1. (c) Theoretical event probabilities and response proportions in blocks of five trials for Cycles 1–2 and Cycles 19–20 on Day 2. (d) Theoretical event probabilities and response proportions in blocks of five trials for Cycles 1–2 and Cycles 19–20 on Day 3.

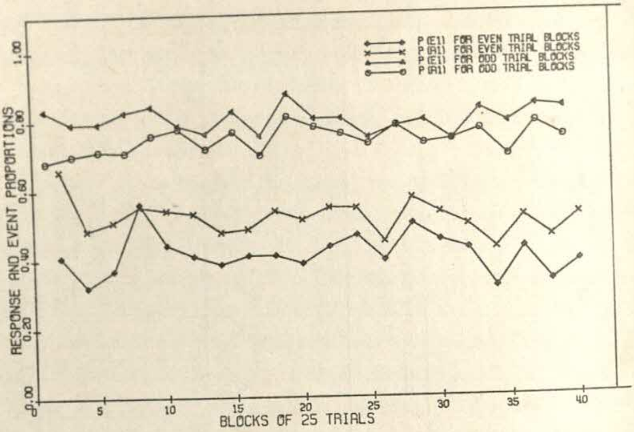


Fig. 2. Average response and event proportions by odd and even 25-trial blocks for Day 1

diction portion of the experiment, the trials following the 1,000 observation trials initiated another cycle of events. The critical test here was whether the subjects would continue to show the cyclic response proportions on blank trials. Figure 1c shows $P(A_1)$ from the initial and final cycles of the test trials compared with the theoretical $P(E_1)$ curve. There was no evidence for cyclic responding without the cyclic event sequence, not even on the earliest blank trials.

However, although the wave pattern degenerated immediately without event feedback, the average $P(A_1)$ continued above .60 for over 300 trials. Figure 3 shows the A_1 response proportions in 25-trial blocks over the full 1,000 test trials. There appears to have been a general trend toward .60 over the 1,000 trials. Average $P(A_1)$ over all trials was .624 and the average over the last 250 trials was exactly .60. Thus, continued feedback does not appear to have been essential for long-term probability matching, replicating results found elsewhere (Reber and Millward, 1968).

Day 3

On Day 3, the subjects again received 1,000 observation trials generated according to the wave schedule, and then they received 1,000 prediction trials with a constant E_1 event probability of .60. Figure 1d shows $P(A_1)$ from the initial and final cycles of the test

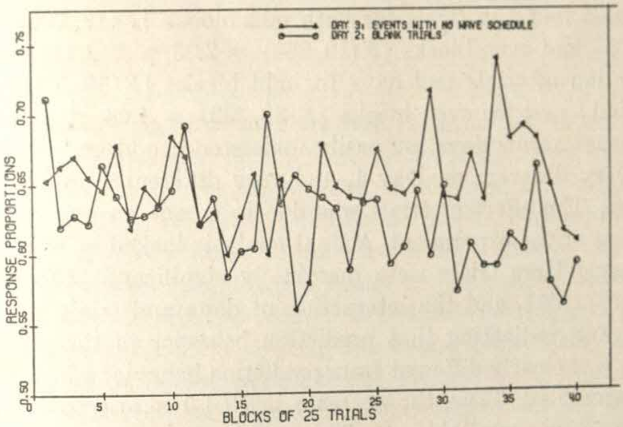


Fig. 3. Response proportions for Days 2 and 3

trials compared with the theoretical curve for $P(E_1)$. Here, as on Day 2, there was no evidence of cyclic responding.

Note, however, that the presence of the events played an important role in controlling the subjects' responses. The triangles in Figure 3 show the average $P(A_1)$ by 25-trial blocks; it did not change over the full 1,000 test trials. The overall average was .65, and the average of the last 250 trials was .66. Hence, the presence of the overt event served to keep response proportions at a level that overshot π, while the absence of the overt event (as on Day 2) produced a gradual shift in response proportion to a level that approximately matched π.

Overall effects

In light of the large systematic effects observed, there seems to be little need to test statistically the reliability of the trial-by-trial tracking of event probabilities. By the gross measure of the proportions of A_1 responses based on 20 blocks of 25 trials, all individual subjects showed some tracking. Further, any analyses involving all three days of the experiment would show significant interactions of trials and days, because of the large deviations from $\bar{\pi}$ on Day 1 but not on Days 2 and 3. Therefore, two analyses were carried out on the response proportions for the odd- and even-numbered trial blocks separately. A highly significant difference in days was found for both odd blocks [$F(2, 28) = 8.96, p < .01$] and even blocks [$F(2, 28) = 15.29, p < .001$]. Less reliable differences were also found in trials for both odd blocks [$F(19, 266) = 1.75$, $p < .05$] and even blocks [$F(19, 266) = 2.03$, $p < .01$], and in the interaction of trials and days for odd blocks [$F(38, 532) = 2.31$, $p < .001$] and for even blocks [$F(38, 532) = 1.44$, $p < .05$]. The differences across days are easily understood: odd and even blocks were very different on Day 1, and very different from the overall average. The effect of trials was due to trends on one or more of the days of the experiment. A final analysis looked at only the last two days. Here trials were marginally significant [$F(39, 546) = 1.46$, $p < .05$], and the interaction of days and trials was highly significant, indicating that prediction behavior in the presence of events was clearly different from prediction behavior when no events were presented. Thus, the decrease in $P(A_1)$ across trials reported for Day 2 was a reliable one. These analyses indicated a significant difference between Days 1, 2, and 3 in the average response proportions. Thus, the three test-trial procedures—wave prediction, blank trials, and prediction of a Bernoulli event sequence—produced different results.

Figure 4 shows the recency curves from Day 1 of the experiment. The curves were very stable and show a change in value only for run lengths of one and two. Since no events were presented on the prediction trials of Day 2, no recency curves exist for those data. Day 3's recency curves were quite flat and are typical of those presented elsewhere (Reber and Millward, 1968; Myers, Gambino, and Jones, 1967).

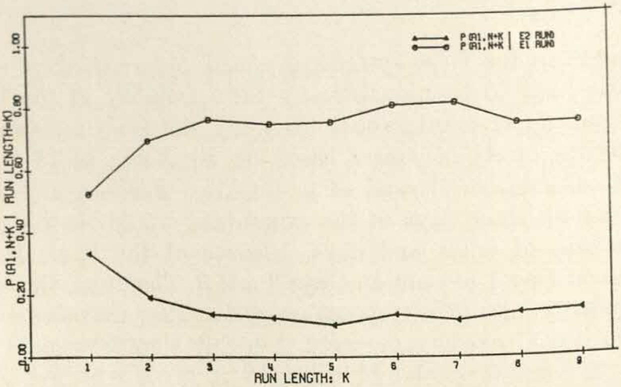


Fig. 4. Recency curves for Day 1

DISCUSSION

The observation technique generally worked as expected: asymptotic response proportions were realized in the initial block of test trials. Thus, as has been reported elsewhere (Reber and Millward, 1968), asymptotes can be approached or achieved in a very short time and without requiring subjects to make overt prediction responses. The effectiveness of the observation trials in producing asymptotic response proportions raises an interesting theoretical problem concerning what the subject is actually learning in a probability-learning experiment. During the observation trials, he may be learning at least two things about the sequence: he may be learning something about the overall proportions of E_1 and E_2 events, or he may be learning something about the wave pattern of the events. The nearly similar asymptotes during all three days of the experiment suggest that learning about the overall $P(E_1)$ was substantial. However, the response proportions on Day 2 indicate that the learning about the wave component was not very impressive. The first few points in Figure 1c show that the subjects' average response sequences quickly degenerated toward an apparently nonstructured sequence.

However, one must be cautious in interpreting these average results. It is possible that individual subjects may have varied their response probabilities but that each individual had a different wave length and amplitude. In such a case, the overall average would tend to be fairly level. This possibility is difficult to check out, since one would have to combine over trials to get reliable estimates of response proportions for individual subjects, and unless one knew the wavelength of each subject's response patterns, any such combination might simply average out the periodicity. If it were possible to be sure that the subject maintained a fixed wavelength for all 1,000 trials, an autocorrelation technique could be used. But if, before the 1,000 trials were completed, the individual wave function damped out, or the wavelength varied, or both, the autocorrelation function would probably be too insensitive to detect periodicity. Because of these difficulties, no statistical proof can be provided for the general observation that individual subjects did not show periodic response patterns.

With respect to overall level of responding, the primary concern is the comparison between our results and those of other studies. The degree of overshooting of $\bar{\pi}$ observed on Day 1 (when the wave pattern was used) was somewhat less than that observed on Day 3 (when a Bernoulli sequence with $\pi = \bar{\pi}$ was used): $P(A_1) = .625$ and .649 respectively. The latter value is more in line with results from other experiments (Lambe, 1968; Pearlmutter, 1968). Overshooting was smaller on Day 1 because the event probabilities were not constant over trials. Note that overshooting is defined as a response proportion greater than the event probability only when the event probability is greater than .5; when $P(E_1) < .5$ with respect to the same response, A_1, overshooting is a response proportion less than the event probability. With the wave schedule used in this experiment, the event proportion was above .5 on only five-eighths of the trials. From previous research, the expected amount of overshooting is about .05; so with the wave schedule the expected amount of overshooting should be about .03. The value observed was .025.

The stability of the overshooting appears to have depended on the presence of the event sequence. When the reinforcing events were removed, during the no-event condition, $P(A_1)$ gradually shifted toward $\bar{\pi}$, while the simple presence of a Bernoulli sequence with $\bar{\pi} = .60$ served to keep subjects' response proportions at an asymptote that overshot $\bar{\pi}$ (see Figure 3). Despite the fact that $P(A_1) = .600$ for the last 250 trials on Day 2, it is possible that the interaction between $P(A_1)$ on Days 2 and 3 was not due to a trend toward $\bar{\pi}$ on Day 2. Instead, the trend may have been toward .5, or there may have been no trend at all, simply a lower overall curve. An attempt to fit the curve by least squares with the equation $P_n = P_\infty - (P_\infty - P_1)(1 - \theta)^{n-1}$ produced $P_1 = .71$, $P_\infty = .625$, and $\theta \approx 1.0$. Thus, from this analysis there is some ambiguity about the assertion that the Day 2 curve was decreasing toward $\bar{\pi}$, although it is fairly clear that it was not heading toward .5.

The predictions that Estes (1957b) derived from statistical learning theory for varying event sequences, such as the wave schedule used here, cannot account for the data. In the present case, the learning curve derived from statistical learning theory has the form

$$P(A_{1,n}) = \bar{\pi} + S(n - 1) - \frac{s}{\theta} + \left[P(A_{1,1}) - \bar{\pi} + \frac{s}{\theta}\right](1 - \theta)^{n-1}$$

$$1 \leq n \leq 13\,,$$

$$P(A_{1,n}) = \bar{\pi} - S(n - 13) + \frac{s}{\theta} + \left[P(A_{1,13}) - \bar{\pi} - \frac{s}{\theta}\right](1 - \theta)^{n-13}$$

$$14 \leq n \leq 38\,, \qquad [2]$$

$$P(A_{1.n}) = \bar{\pi} + S(n - 38) - \frac{s}{\theta} + \left[P(A_{1.38}) - \bar{\pi} + \frac{s}{\theta}\right](1 - \theta)^{n-38}$$

$$39 \leq n \leq 50\,.$$

The parameters S and $\bar{\pi}$ were defined in Equation 1 and depend on the reinforcing-event schedule. The single parameter to be estimated is the learning parameter, θ. Equation 2 repeats every 50 trials; that is, after the first 50 trials, $P(A_{1.51})$ replaces $P(A_{1.1})$ in the first of the three expressions. Because of the way the event schedule was generated, namely, with $P(E_{50}) = P(E_{51}) = \bar{\pi}$, the value of $P(A_{1.51})$ must be determined from $P(A_{1.50})$ by the equation

$$P(A_{1.51}) = (1 - \theta)P(A_{1.50}) + \theta\bar{\pi}\,. \qquad [3]$$

It can be shown that if these equations are averaged for the first 25 trials (odd blocks) and then for the second 25 trials (even blocks), the averages approach constants as n approaches ∞. This particular prediction relates to the observation made about Figure 2: that the response proportions approach constants except for fluctuation due to the event variations. The expected averages could be used to estimate the learning parameter if it were not for the fact that overshooting was observed. The average of the predicted odd and even blocks must be $\bar{\pi}$ according to Equation 2; if it is not, the equation cannot be used to estimate θ.

However, there is a more serious discrepancy between statistical learning theory and the data than overshooting. The discrepancy is that at asymptote the peak of the response proportions occurs at the peak of the event proportions. It can be shown that the response proportion should theoretically peak at the trial on which the curve of response proportion crosses that of event proportions. A large learning rate (θ) increases the amplitude of the theoretical response-proportion function, causing it to cross the event function at an earlier trial. However, the only way to have the two curves peak together would be to set θ at 1.0, in which case the curves would have the same amplitude.

In order for a theory to predict simultaneous peaking of the event sequence and the response proportion, it must contain some operation that averages or integrates over past trials. A theoretical possibility that has recently adequately described some probability-learning data is the weighted-average model (Friedman, Carterette, and Anderson, 1968). The basic assumption of the model is that the subject's response proportion on Trial n, p_n, is a function of the averaged weights associated with past trials. The subject is viewed as an information integrater, and the events as information, not reinforcers.

Let e_n be a random variable equal to 1 with $P(E_{1,n})$ and equal to 0 with $P(E_{2,n})$. Then,

$$P_n = w_0 l_0 + \sum_{i=1}^{k} w_i e_{n-i} , \qquad [4]$$

where the w_i values represent weights associated with past events and $\Sigma_{i=0}^{k} w_i = 1.0$. Application of the model for $k = 1$, 2, and 3 led to nearly equally good fits. In all cases, the estimation technique was by a hill-climbing procedure using a simple least-squares minimization on the average 50-trial wave function illustrated in Figure 1b. For all three k values, $l_0 = .656$ and $w_0 = .412$. For $k = 2$, $w_1 = .378$, and $w_2 = .210$, while for $k = 3$, $w_1 = .459$, $w_2 = .055$, and $w_3 = .085$. The three parameters w_1–w_3 were fairly unstable in that various values gave nearly identically good fits. In most cases, however, the values were near those listed above. Instability of these parameters is not unreasonable, considering that they simply weight the probabilities of events on Trials $n-1$, $n-2$, and $n-3$, which are very similar in value. The low values of w_2 and w_3 illustrate the relatively weak influence of events more than two trials back. Figure 5 presents the tracking performance as predicted by the $k = 2$ case. The fit is essentially perfect. The conclusion that subjects were anticipating the shifting wave form is not necessary with the information-averaging interpretation provided by the model of Friedman et al.

Other aspects of the averaging model are important here. One is that the model does not predict recency curves that increase monotonically to 1.0 but, rather, recency curves for E_2 runs that have an asymptote at $w_0 l_0$ and recency curves for E_1 runs that have an asymptote at $1 - w_0(1 - l_0)$. The predictions using the above parameters for $k = 2$ are .270 and .858, which are both too large (see Figure 4) but at least reasonably close compared to the .0 and 1.0 predictions of statistical learning theory. Also, note the fact that l_0 seems to reflect the asymptotic response proportion. Finally, if an odd block of 25 trials has, on the average, E_1 events with probability .8 and an even block has E_1 events with probability .4, then the predictions for these blocks are .740 and .505. These

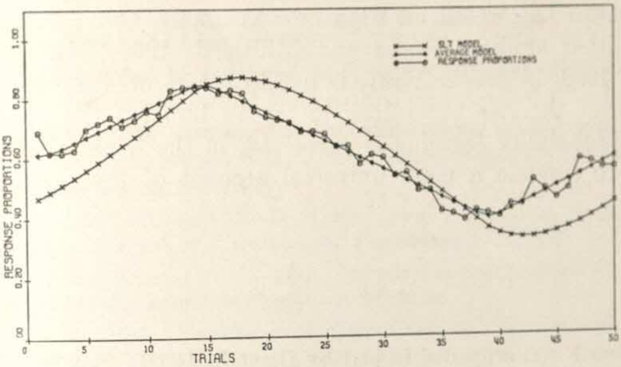


Fig. 5. Response proportions as predicted by statistical learning theory (SLT) and by the averaging model

values agree nicely with the .739 and .516 levels of response proportions presented in Figure 2.

However, there are serious questions raised by the averaging model. First, although the model has a large advantage over many other models in that it has two or three parameters, it probably gains its power of prediction by incorporating the psychological fact that past events play an important role in determining response probability. It is important to note that the model treats these events as components of information rather than discrete reinforcers, which may be an important innovation in theorizing about probability learning. Another question concerns the determination of weights, w_i. If the averaging model does not specify the mechanism for altering the weights, then the successful predictions reported here are little more than curve fitting. If weights change systematically with the event schedule, then psychological theory requires a rule or a mechanism by which the weights are related to the event sequence.

At this point, we feel, no adequate theory of probability learning yet exists. Conditioning theories, such as statistical learning theory, are in error in so many respects that the existence of a simple correcting axiom seems unlikely. None of the hypothetical models has reached a state of development that allows any adequate evaluation—they say very little about most of the facts known quantitatively about probability learning. Scattered throughout the literature are ad hoc models that account for specific results fairly well, but none has the necessary breadth to account for the wealth of data now available. Our purpose here, as in other studies (Reber and Millward, 1968; Millward and Reber, 1968, in preparation), is not to credit or discredit any of the extant theories but rather to try to broaden the focus with which probability learning is perceived, in the hope of a restructuring to provide a more universal account of the phenomenon.

Notes

This research was supported in part by Grant MH-11255 from the National Institutes of Mental Health to Richard B. Millward and by Grant APA-240 from the National Research Council of Canada to Arthur S. Reber, who is now at Brooklyn College, City University of New York, Brooklyn, New York 11210. Received for publication May 20, 1970.

1. The computer program for generating the event sequences was written by Thomas Wickens. Partial support for computing was received from Grant GP-4825 from the National Science Foundation to the Brown University Computing Center.

References

Anderson, N. H. 1969. Effect of first-order conditional probability in a two-choice learning situation. *Journal of Experimental Psychology* 59:73–93.

Anderson, N. H. 1964. An evaluation of stimulus sampling theory: Comments on Professor Estes' paper. In *Categories of human learning*, ed. A. W. Melton. New York: Academic Press.

Estes, W. K. 1957a. Of models and men. *American Psychologist* 12:609–617.

Estes, W. K. 1957b. Theory of learning with constant, variable, or contingent probabilities of reinforcement. *Psychometrika* 22:113–132.

Friedman, M. P., Carterette, E. C., and Anderson, N. H. 1968. Long-term probability learning with random schedule of reinforcement. *Journal of Experimental Psychology* 78:442–455.

Greeno, J. G., and LaBerge, D. 1963. Sequential dependencies and nonreinforcement in probability learning. *Journal of Experimental Psychology* 66:547–552.

Lambe, R. A. 1968. Human binary prediction under two schedules of response-contingent reinforcement. Unpublished doctoral dissertation, Brown University.

Millward, R. B. 1967. Two learning models for two-choice conditioning experiments involving nonreinforced trials. *Psychological Reports* 20:1211–1229.

Millward, R. B., and Reber, A. S. 1968. Event-recall in probability learning. *Journal of Verbal Learning and Verbal Behavior* 7:980–989.

Millward, R. B., and Reber, A. S. In preparation. Probability learning: Contingent event schedules with lag.

Myers, J. L., Gambino, B., and Jones, M. R. 1967. Response speeds in probability learning. *Journal of Mathematical Psychology* 4:473–488.

Pearlmutter, F. 1968. Investigation and implications of the noncontingent success and observation techniques in human probability learning. Honors thesis, Brown University.

Reber, A. S., and Millward, R. B. 1968. Event observation in probability learning. *Journal of Experimental Psychology* 77:317–327.

Vitz, P. C., and Hazan, D. N. 1969. Memory during probability learning. *Journal of Experimental Psychology* 80:52–58.

Young, F. W. 1967. Periodic probability learning. *Psychonomic Science* 8:71–72.

PROBLEM SHIFTS AND HYPOTHESIS BEHAVIOR IN CONCEPT IDENTIFICATION

James R. Erickson
Ohio State University

When college-age subjects were given thorough instructions stressing the nature of the solution of the problems and pointing out stimulus dimensions, a reversal shift was more difficult than an extradimensional shift, or there was no difference between them. When subjects were given brief instructions that they were to learn which stimuli fell into each of two categories, the reversal shift was easier. A hypothesis-sampling model accounts for the first set of data; Kendler and Kendler's mediational theory of concept learning accounts for the second.

A very common finding in the literature on concept shifts is that college students in two-choice tasks find a reversal shift easier to learn than an extradimensional shift (for a review, see Wolff, 1967). Occasionally, however, no difference between a reversal and an extradimensional shift has been found in situations where subjects receive more thorough instructions than usual about the dimensions along which the stimuli vary or about the nature of the solution of the problems (Johnson, 1966; Johnson, Fishkin, and Bourne, 1966; Schvaneveldt and Kroll, 1968).

The studies showing a reversal shift to be easier than an extradimensional shift have been widely quoted and have received a great deal of theoretical attention, usually in support of mediational theories of concept learning (Kendler and Kendler, 1962, 1968). Although hypothesis-sampling theories of concept learning have not dealt with these data in any detail, the no-memory models of Bower and Trabasso (1964) and of Restle (1962) predict no difference between the shift types,[1] while some models assuming short-term memory of previous stimuli or of hypotheses tested predict that a reversal shift should be more difficult than an extradimensional shift, rather than the reverse (Erickson, 1968; Levine, 1969; Trabasso and Bower, 1966).

In a recent study testing various assumptions about hypothesis sampling in concept identification, Erickson, Block, and Rulon (1970) found that a reversal shift *was* more difficult than an extradimensional shift. However, there were possible artifacts in that study which might have influenced the results. This paper presents the results of five studies undertaken in an attempt to determine the conditions under which a reversal shift is more difficult than an extradimensional shift for college-age subjects.

EXPERIMENT I

The major purpose of Experiment I was to replicate and extend the results of Erickson et al. (1970) and to further examine hypothesis-sampling behavior in concept-identification problems.

Method

-Subjects-The subjects were 128 students from the introductory-psychology subject pool. An additional 31 subjects were run but replaced when they failed to learn one of their problems within the time period (60 min) or because of apparatus failure.

-Apparatus and materials-The stimuli were square figures varying along four binary dimensions: single or double, solid or broken border, a line inside which was either horizontal or vertical, in either the upper or lower half of the square. The stimuli were rear-projected onto a screen at the front of the subject's booth. The subjects sat at a table with a small response box containing two buttons and two feedback lights, labeled A and B. They wore earphones into which low-level white noise was projected to mask distracting noise from the projector, timer, relays, and so forth.

-Design-The subjects were randomly assigned to the eight conditions formed by crossing two types of shift (reversal and extradimensional) with four preshift criterion levels (2, 4, 6, or 10 consecutive three-trial blocks on which the correct hypothesis was used). There were 16 subjects per condition. The preshift problem was chosen randomly for each subject, with the restriction that each of the eight possible solutions be used twice within each condition. For the reversal-shift subjects, the shift problem was a reversal of the preshift problem; for the extradimensional-shift subjects, the shift problem was randomly chosen from one of the two dissimilar dimensions with the restriction that each dimension be used an approximately equal number of times. For example, if the relevant dimension for the preshift problem was one of the border dimensions, the extradimensional shift was to one of the interior dimensions. The dimension was chosen randomly ahead of time, but the particular response assignment within the dimension was chosen at the time of the shift so that all subjects received 'error' feedback on the first shift trial. The shift-problem criterion was six consecutive three-trial blocks on which the correct hypothesis was used.

-Procedure-Each subject was given detailed instructions about the nature of the solution of the problems, instructions stressing that the experimenter would classify each stimulus A or B according to the value of one of the dimensions. The subjects then solved two practice problems involving two binary dimensions; before solving each problem the subjects were shown examples of the stimuli to be used and the various dimensions were pointed out. For the final problem a partial-feedback, hypothesis-monitoring technique was used (Erickson, 1968). The subjects were told that they would not receive feedback on every trial and that when no feedback was given they were to continue responding as if they had been correct.

With four binary dimensions 16 stimuli are possible. Six sets of 15 of these stimuli were formed; each set was divided into five three-trial blocks, such that each block permitted the inference of a hypothesis. Orderings of the six sets were counterbalanced across subjects within conditions as much as possible. Feedback was given every third trial. After the preshift problem was solved to the appropriate criterion, the subject was shifted without warning to his shift problem. On trials when feedback was given, it was given 2 sec after the response and remained on for 2 sec, after which both the feedback light and slide went off. Then 10 sec later the next slide came on. When no feedback was given, the slide went off 4 sec after the response, followed by the 10-sec intertrial interval.

Results

The major results of Experiment I are shown in Figure 1, which shows the number of three-trial blocks to the start of the criterion run for the various groups in the experiment. It is clear that the reversal-shift and extradimensional-shift groups were equivalent on the preshift task but that the reversal-shift groups found the shift task more difficult.

For the preshift problem, only the effect of preshift criterion was significant [$F(3, 120) = 2.77$, $p < .05$]. As the criterion became more stringent, subjects took longer to reach it. The reason for this is probably the fairly high win-shift probability reported below; subjects with higher criterion levels sometimes shifted away from the solution hypotheses during part of what would have been the criterion run, then returned to it immediately. This could be because of what Levine has called an 'oops' error, when the subject accidentally makes a wrong response. Because the criterion used in this study required a series of successive correct hypotheses, and any 'oops' error would be indicated as a change in hypotheses, the

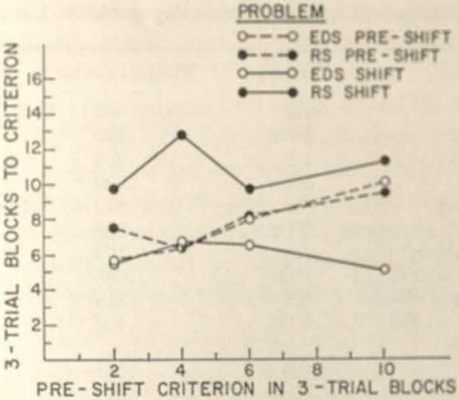


Fig. 1. Number of three-trial blocks (feedback trials) to criterion as a function of preshift criterion and shift type (extradimensional, EDS; reversal, RS) for preshift and shift problems

occurrence of these errors would have a larger effect on groups with a longer criterion.

Analysis of the shift-problem data showed that the effect of shift type was highly significant [$F(1, 120) = 20.94$, $p < .001$]. The effect of preshift criterion and the interaction of criterion and shift type were not significant [($F(3, 120) = .73$ and .46 respectively].

Some summary hypothesis-sampling statistics are presented in Table 1.[2] The win-shift probability (the probability of changing hypotheses after a 'correct' feedback) was rather high. Comparable probabilities were found by Erickson et al. (1970), but other studies have reported much lower ones.

The p(lose-stay), the probability of keeping the same hypothesis after receiving an 'error' feedback, was fairly low, about the level reported in other studies. The p(local consistency), the probability that a new hypothesis is chosen which is consistent with the information on the previous feedback trial, was also about the level reported previously. Several models have been presented which postulate local consistency; the reported empirical p(local consistency) indicates that the models cannot be correct as formulated.

The p(reversal hypothesis), the probability that a new hypothesis is the reversal of the prior one, was quite low, below chance. Of the hypothesis reversals that did occur, many of them came on the

Table 1. Summary of hypothesis-sampling statistics, Experiment I

	Shift type	Preshift problem	Shift problem
p (win-shift)	RS	.366	.299
	EDS	.430	.314
p (lose-stay)	RS	.058	.070
	EDS	.039	.142
p (local consistency)	RS	.872	.824
	EDS	.835	.874
p (reversal hypothesis)	RS	.120	.099
chance = 1/7 = .143	EDS	.095	.042
p (similar-EDS hypothesis)	RS	.319	.379
chance = 2/7 = .286	EDS	.403	.393
p (dissimilar-EDS hypothesis)	RS	.561	.522
chance = 4/7 = .571	EDS	.502	.565

Note: RS is reversal shift, EDS is extradimensional shift.

first hypothesis shift. After that they were fairly rare. New hypotheses tended to be on different dimensions than the prior ones. This is apparently the main reason a reversal shift was more difficult than an extradimensional shift (EDS).

The *p*(similar-EDS hypothesis) is the probability that a new hypothesis is chosen on the similar dimension (e.g., a shift from one border dimension to the other). This probability is of interest because of Nelson's results (1969) showing that an extradimensional shift to a dissimilar dimension was easier than one to a similar dimension. Since all extradimensional shifts in the present study were to a dissimilar dimension, this result could not be directly verified here. However, it should be noted that *p*(similar-EDS hypothesis) was greater than chance in all cases. This would not be predicted from Nelson's results. The *p*(dissimilar-EDS hypothesis) was close to chance in all cases.

EXPERIMENTS II-IV

The data from Experiment I fully replicate the findings of Erickson et al. (1970). They show that a reversal shift is sometimes more difficult than an extradimensional for college-age subjects and that the relative difficulty does not seem to be a function of preshift criterion. The next three experiments provide further replication of the results.

Experiment II was a partial replication of Experiment I, except that the subjects were given feedback on the correct classification of the stimulus on every trial, the procedure normally used in concept-learning experiments. It seemed possible that the partial-feedback technique, per se, might have influenced the results.

In Experiments I and II, the stimuli varied along four binary dimensions. Since most of the earlier studies that found a reversal shift easier than an extradimensional shift had used stimuli varying along only two binary dimensions, Experiments III and IV then used two-dimensional stimuli. In Experiment III, the subjects first solved a practice problem; in Experiment IV no practice problem was used.

Method

-Subjects-The subjects in the three studies were students from the introductory-psychology subject pool. There were 100 subjects in Experiment II, 48 in Experiment III, and 40 in Experiment IV. Twelve additional subjects were run but replaced because they did not solve one of their problems or because of apparatus failure.

-Apparatus and materials-The apparatus was the same as in Experiment I. The stimuli for Experiment II were the same as in Experiment I, while the stimuli for Experiments III and IV used two of the same dimensions: solid or broken border, and the inside line in either the upper or the lower half of the square.

-Design and procedure-In Experiment II, 25 subjects were randomly assigned to each of the four conditions formed by crossing shift type (reversal or extradimensional) with preshift criterion (6 or 10 consecutive correct responses). In Experiments III and IV, the subjects were randomly assigned to either a reversal-shift group or an extradimensional-shift group, with equal numbers in each. The preshift criterion in the latter two studies was 10 consecutive correct responses; this was the postshift criterion in all three studies. Details of timing conditions and of counterbalancing of preshift solution and shift solution were the same in all three studies as in Experiment I.

The general procedure was the same as in Experiment I, except for the practice problems. All subjects received thorough instructions as to type of solution and the dimensions involved in all problems. In Experiment II the subjects solved two two-dimensional practice problems before their main problem. In Experiment III they solved a four-dimensional practice problem before solving their main problem; and in Experiment IV no practice problem was used.

Results

-Experiment II-There were no significant effects due to the preshift criterion (all F ratios < 1.0), so only the effects of shift type are presented, averaged across preshift criterion. On both the preshift problem and the shift problem, the reversal-shift groups made more errors before reaching criterion. The mean errors on the preshift problem were 3.00 for the reversal-shift subjects and 2.02 for the extradimensional-shift subjects [$F(1, 96) = 4.59$, $p < .05$]. The mean errors on the shift problem were 5.14 and 3.50 for the reversal-shift and extradimensional-shift subjects respectively [$F(1, 96) = 3.22$, $p < .07$]. Although the results are somewhat ambiguous because of the preshift-problem difference, they are in the same direction as those of Experiment I and of Erickson et al. (1970).

–Experiment III–There were no significant differences between reversal-shift and extradimensional-shift groups on any measure. The mean numbers of errors for the reversal-shift group on the practice, preshift, and shift problems were 5.96, 1.13 and 2.04 respectively, while for the extradimensional-shift group the corresponding means were 5.58, .92, and 1.71 (all t values < 1.0). While the reversal shift was numerically slightly more difficult than the extradimensional shift, the difference was clearly small and unimportant.

–Experiment IV–The mean number of errors for the preshift problem was 1.10 for both reversal-shift and extradimensional-shift groups. The mean number of errors on the shift problem was 2.80 for the reversal-shift group, and 1.60 for the extradimensional shift group [$t(38) = 3.33$, $p < .01$].

In Experiment IV, where no practice problem was used, the reversal shift was again more difficult than the extradimensional shift, while in Experiment III, where a practice problem was used, there was no difference between the two shift types. A supplementary analysis of variance indicated that the interaction between experiments and shift type was significant [$F(1, 84) = 4.62$, $p < .05$]. While the reason for the different results in the two studies is not clear, it has been demonstrated that the relative difficulty of reversal and extradimensional shifts was not due to the complexity of the stimuli used or to the partial-feedback technique.

EXPERIMENT V

Experiments I–IV all used thorough instructions, pointing out explicitly the dimensions along which the stimuli would vary and stressing that the experimenter would choose one of the dimensions and label each stimulus A or B according to the value of that dimension. As was noted in the introduction, the previous studies that found no difference between a reversal and an extradimensional shift also used thorough instructions—although it should be noted that none of them found a reversal shift to be more difficult than an extradimensional shift. The studies that found a reversal shift to be relatively easy have, for the most part, used rather sketchy instructions, telling the subject only that the experimenter would label the stimuli A or B (or some other binary classification) and that it was the subject's task to determine how the stimuli were labeled.

In Experiment V, the effect of thorough versus brief instructions was examined explicitly, with half of the subjects receiving each type followed by either a reversal-shift or an extradimensional-shift problem. The general design was similar to that used by Johnson et al. (1966), with stimuli varying along two binary dimensions and no practice problem.

Method

-Subjects-The subjects were 80 students in the introductory-psychology subject pool. Seven additional subjects were run but did not solve one of their problems.

-Apparatus and materials-These were the same as in Experiments III and IV.

-Design and procedure-The subjects were randomly assigned in equal numbers to the four groups formed by crossing instruction type (thorough or brief) and shift type (reversal or extradimensional). The subjects receiving thorough instructions received the same instructions that were used in Experiments I-IV. Those receiving brief instructions were told only that the stimuli would be systematically labeled A or B and that their task was to determine which figures were labeled A and which B. No practice problems were used. Details of counterbalancing and timing conditions were the same as in the earlier studies.

Results

The results of Experiment V are given in Table 2, which shows the mean number of errors to solution in the various conditions. Analysis of the preshift-problem data showed that neither of the main effects nor their interaction was significant. In the analysis of the shift-problem data, neither the main effect of shift type nor that of instructions was significant [$F(1, 76) = .41$ and 2.55 respectively]. The interaction of shift type and instructions, however, was highly significant [$F(1, 76) = 8.73, p < .01$]. The meaning of the interaction is clear and unmistakable; it is the classic 'crossover' interaction. The subjects who received thorough instructions found the reversal shift more difficult than the extradimensional shift, while the subjects who received the brief instructions found the reversal shift easier than the extradimensional shift.

Table 2. Mean number of errors on preshift and shift problems as a function of instructions and type of shift, Experiment V

Type of instructions	Shift type	Preshift problem	Shift problem
Thorough	RS	1.75	4.05
	EDS	2.35	2.60
Brief	RS	1.85	1.20
	EDS	1.20	3.45

Note: RS is reversal shift, EDS is extradimensional shift.

The effect of the thoroughness of the instructions verifies the generalization drawn from our review of the literature. When subjects are given brief instructions telling them only that the stimuli must be classified somehow into two groups, a reversal shift is easier than an extradimensional shift; when they are told the type of rule to be used and the dimensions along which the stimuli will vary, the reversal shift is sometimes more difficult than the extradimensional shift, and sometimes there is no difference, presumably depending on the care taken to ensure that the instructions are understood.

DISCUSSION

It is apparent that the process by which subjects solve concept-identification problems differs as a function of instructions. For thoroughly instructed subjects, the process seems very much like that inferred from studies of hypothesis sampling. Those models assume that the subject solves concept-identification problems by sampling hypotheses from a pool of possible solutions to the problem, using whatever prior information he can remember when choosing new hypotheses—using a combination of prior hypotheses tried-but-rejected and prior stimulus information, especially from the immediately preceding feedback trial. The present data show that when the subject chooses a new dimension (or dimensions) to test, it is different from the one his last hypothesis was based on and that usually, but not always, the new hypothesis is consistent with information from the previous trial.

This type of model—assuming that subjects choose new hypotheses based on a different dimension than the old hypothesis, and assuming that they remember and do not resample previously rejected hypotheses for a period of time—predicts that a reversal shift should be more difficult than an extradimensional shift, the result found in Experiments I, II, IV, and, when thorough instructions were given, in Experiment V. Since these models were developed primarily to account for hypothesis-sampling behavior in simple concept-identification experiments, it is of some interest that they can also account for behavior in problem-shift experiments (when subjects are thoroughly familiar with the demands of the situation).

When subjects are given brief instructions telling them only that they have to learn to sort stimuli into two classes, they almost certainly do not sample hypotheses in the above manner. Their task, in fact, is one of complete learning (Haygood and Bourne, 1965), rather than of attribute learning or concept identification, in that they have to learn the conceptual rule in addition to learning the relevant dimension. It would not be expected that the same model would describe hypothesis sampling for both complete-learning and the concept-identification tasks. Furthermore, it may not be the case that the subjects solve a complete-learning task by sampling hypotheses at all. In addition, it seems likely that subjects sometimes solve their preshift problems under brief instructions without having full knowledge as to all the dimensions along which the stimuli vary—that they solve the problems having learned mainly the relevant dimension. This could be another factor causing a reversal shift to be easier than an extradimensional shift for briefly instructed subjects. Data on this point would be quite valuable but have not yet been reported.

The processes of attention and/or mediation discussed by the Kendlers (1962, 1968) and others are able to account for behavior under brief instructions, but not under thorough instructions. Whether or not mediational theory can be revised to account for these data remains to be seen. At the present time there seem to be three levels of behavior possible: a single-stage level, a two-stage mediation level, and a hypothesis-sampling level, with only the mediational level producing faster shifts under reversal-shift conditions than under extradimensional-shift conditions. It would be valuable to know whether the ability to perform at the hypothesis-testing level is a function of developmental stage or is solely a function of the ability of the experimenter to communicate the demands of the problem to the subject.

Notes

This research was carried out at the Human Performance Center and was supported in part by Grant MN-12113-03 from the National Institutes of Health. Portions of the data were presented in a paper at the meetings of the Psychonomic Society in Saint Louis, Missouri, 1968. The author is indebted to Jean Butler, who was responsible for data collection and portions of the analyses.

1. Strictly speaking, the prediction is no difference following the first error after the shift has occurred.

2. Previous studies in which these statistics have been reported include Erickson (1968), Erickson and Block (1969), Erickson et al. (1970), and Levine (1969).

References

Bower, G., and Trabasso, T. 1964. Concept identification. In *Studies in mathematical psychology,* ed. R. C. Atkinson, pp. 32–96. Stanford, Calif.: Stanford University Press.

Erickson, J. R. 1968. Hypothesis sampling in concept identification. *Journal of Experimental Psychology* 76:12–18.

Erickson, J. R., and Block, K. K. 1969. Concept learning with reinforcement of hypotheses and misinformative stimulus labeling. *Psychonomic Science* 14:183–184.

Erickson, J. R., Block, K. K., and Rulon, M. J. 1970. Some characteristics of hypothesis sampling in concept identification. *Psychonomic Science* 20:103–105.

Haygood, R. C., and Bourne, L. E., Jr. 1965. Attribute and rule learning aspects of conceptual behavior. *Psychological Review* 72:175–195.

Johnson, P. J. 1966. Factors affecting transfer in concept identification problems. *Journal of Experimental Psychology* 72:655–660.

Johnson, P. J., Fishkin, A., and Bourne, L. E., Jr. 1966. Effects of procedural variables upon reversal and interdimensional shift performance, II. *Psychonomic Science* 4:69–70.

Kendler, H. H., and Kendler, T. S. 1962. Vertical and horizontal processes in problem solving. *Psychological Review* 69:1–16.

Kendler, H. H., and Kendler, T. S. 1968. Mediation and concept behavior. In *Psychology of learning and motivation: Advances in research and theory,* vol. 2, ed. K. W. Spence and J. T. Spence, pp. 197–244. New York: Academic Press.

Levine, M. 1969. Neo-noncontinuity theory. In *Psychology of learning and motivation: Advances in research and theory,* vol. 3, ed. G. Bower and J. T. Spence, pp. 101–134. New York: Academic Press.

Nelson, T. O. 1969. Dimension similarity in concept identification and extradimensional shifts. *Journal of Experimental Psychology* 82:143–147.

Restle, F. 1962. Selection of strategies in cue learning. *Psychological Review* 69:329–343.

Schvaneveldt, R. W., and Kroll, N. E. A. 1968. Effects of cue associations in concept shifts. *Journal of Verbal Learning and Verbal Behavior* 7:474–478.

Trabasso, T., and Bower, G. 1966. Presolution dimension shifts in concept identification: A test of the cue-sampling axiom. *Journal of Mathematical Psychology* 3:163–173.

Wolff, J. L. 1967. Concept-shift and discrimination-reversal learning in humans. *Psychological Bulletin* 69:369–408.

UNIT-SEQUENCE INTERFERENCE AND SHORT-TERM RECALL

Thomas W. Turnage and Jenny I. Steinmetz
University of Maryland

Two experiments were performed to evaluate unit-sequence interference in short-term serial recall. Both experiments replicated the previous finding of a curvilinear relationship between item frequency and recall, a relationship predicted by interference theory. However, since the data showed little direct evidence for associative disruption of serial order, an alternative interpretation that does not stress concepts of associative chaining was considered.

The hypothesis of unit-sequence interference (Underwood and Postman, 1960) predicts that disruption of serial recall should increase with the linguistic frequency of the verbal units involved. This disruption is assumed to involve preexperimental associations (A–B) that interfere with experimentally required ones (A–C). Unfortunately, clear-cut empirical evidence on this aspect of interference theory has not been particularly impressive (Slamecka, 1966; Underwood and Ekstrand, 1966), partly because the test situations often involve both interference and facilitation from A–B sources (see Postman, 1967). However, there are experimental situations involving short-term recall in which frequency of verbal items and serial recall seem related in a curvilinear fashion as might be expected from interference theory (Turnage and McCullough, 1968). Therefore, the focus of the present two studies was on the role of unit-sequence interference in such short-term recall situations, since its effects seem clearly emphasized there.

EXPERIMENT I

Two major questions were asked here. First, would a free-association task involving stimuli that were to be recalled serially *increase* unit-sequence interference? Second, would familiarization training with these stimuli *reduce* unit-sequence interference? Presumably, the association task would 'prime' the pool of A–B associates, thereby increasing the opportunity for A–B, A–C competition; while familiarization training might provide an early opportunity for discriminating between A–B and A–C competitors, thereby reducing interference during serial recall. Both tasks were expected to attenuate letter-sequence interference. Such interference is reflected in lack of item availability and in spelling errors, and varies inversely with item frequency (Underwood and Postman, 1960).

Method

-Materials–Four sets of 20 two-syllable items were used to prepare the serial-recall lists. The median G-count values (Thorndike and Lorge, 1944) for these sets were 0, 24, 40, and 125. (All A values in the G count were assumed to equal 75, while AA values were assumed to equal 125.) These sets were selected from a pool of items described by Turnage and McCullough (1968). The items in the 0-frequency set were paralogs, and the remaining items were nouns. Items within a set were low in interitem associative strength (IAS). That is, they did not tend to call each other out as associatives, according to norms.

Two serial lists of 12 items each were constructed for each set of items, giving a total of eight different serial lists. At each frequency level, these lists had 4 items in common. A random serial order was assigned to each list, and this order was designated to be the order in which the lists were presented for recall.

-Design and procedure–The experimental lists were presented by a slide projector placed behind a viewing screen. For the recall task, each item was presented at a 1-sec rate. The subjects were instructed that their task was to remember the items being presented, in a backward order. That is, they were to write down the last item in the list first, the next to the last item second, and so on (backward recall, unpaced).

For one-third of the subjects, the recall task was their only task (control groups). For another third of the subjects, the recall task was preceded by a free-association task involving the items to be in the recall task (association groups). In this free-association task, each item was presented at a 30-sec rate. Although only 30 sec were allowed for associating to each item, the instructions and procedure were quite similar to Noble's (1952). The order of stimulus presentation on the free-association task was identical to that for the recall task, which followed immediately.

The remaining subjects received four familiarization trials involving four random orders of the list of items that they were to recall subsequently (familiarization groups). In this condition, the subjects were told to watch the screen but not to write down any of the items, which they would recall later. Items were presented at a 1-sec rate. The recall task followed immediately. Due to a procedural error, the order of item presentation for learning was different from that for the last familiarization trial. However, this did not appear to produce a serious bias in subsequent comparisons among groups.

–Subjects–A total of 96 subjects who were meeting the requirements of an introductory course was used. As they entered the laboratory, they were assigned in groups of 4 to each of the 24 experimental conditions: 3 conditions of prior testing × 4 frequency levels × 2 experimental lists. Assignments were made on the basis of an unsystematic, predetermined testing order.

Results

Table 1 shows the average number of *pair* scores as a function of the various experimental conditions. A pair score reflects the correct serial order of two items with respect to each other, regardless of preceding or subsequent errors, and in this sense is similar to typical measures of serial learning. The maximum pair score that may be obtained in a list of n items is $n - 1$. An item had to be spelled correctly in every detail to enter into a pair score.

Table 1. Mean pair score as a function of conditions of prior testing and frequency, Experiment I

	List frequency				
	0	24	40	125	Row means
Association groups	1.38	1.50	2.25	1.50	1.66
Familiarization groups	1.00	1.62	2.12	2.50	1.81
Control groups	.25	.88	1.88	1.25	1.06
Column means	.88	1.33	2.08	1.75	1.51

Analysis of the pair scores for Table 1 showed the main effects of groups [$F(2, 84) = 3.16$] and of frequency [$F(3, 84) = 4.22$] to be significant at the .05 level. However, the F ratio for their interaction was less than unity. Both the linear [$F(1, 84) = 4.80, p < .05$] and quadratic [$F\ (1, 84) = 6.06, p < .05$] trends relating frequency and pair scores were reliable, according to an analysis for unequal intervals. Thus, these data replicate the curvilinear relationship, predicted from interference theory, that was previously found for pair scores in short-term recall (Turnage and McCullough, 1968), but they do not support the present experimental hypotheses. A Duncan range test for group means indicated that averaged over frequency levels, the association and familiarization groups differed from the control groups, but not from each other, at the .05 level (see Table 1).

EXPERIMENT II

The experimental focus here, in contrast to Experiment I, was on disruption of A–B rather than A–C associates. The analysis of such disruptions was conceptualized in terms of the A–B, A–C interference paradigm. That is, changes in preexperimental associations (A–B) following serial learning and recall of experimental associations (A–C) were to be evaluated by means of an association task.

Method

-Materials-Three sets of 24 two-syllable items were used to prepare the serial-recall lists. These items came from the same pool as was used in Experiment I and thus were not highly associated with each other within sets, according to norms. The median G-count values (Thorndike and Lorge, 1944) of the three sets were 0, 12, and 75. Two lists of 12 items each were constructed from each set by arbitrarily dividing the sets in half. Random serial orders were assigned to each of these 6 lists. Then, these orders were reversed, giving a total of 12 different lists with 4 at each frequency level.

-Design and procedure-The design involved two degrees of 'familiarization' training (one or four study trials), two recall tasks (backward or forward), and three levels of list frequency (high, medium, or low). The backward-recall task was identical to that employed in Experiment I, while the forward-recall task required serial reproduction of experimental items in the same order in which they were presented. The forward-recall task was included to see if prior findings might be specific to backward-recall data. In both cases, 30 sec were allowed for the subject to write his responses (unpaced serial recall). For experimental subjects, items were presented one *or* four times (in the *same* serial order) at a 1-sec rate before recall.

Immediately after the recall test, the experimental subjects were given a free-association task involving the 12 recall stimuli. These subjects had 30-sec to associate to each stimulus, as did the control subjects, who had not participated in the study-recall task. The controls provided a baseline against which to evaluate changes in associations as a function of prior serial learning. It should be especially noted that the free-association instructions specified that *both* B and C associates were appropriate—that items from the list (C) or from off the list (B) could be given. Thus, the requirements of the task could be considered analogous to those for the method of modified free recall (MMFR). Presentation of all experimental materials was by slide projector placed behind a viewing screen.

–Subjects–A total of 192 subjects was used for the experimental conditions, and a total of 48 subjects was used for the control conditions. All the subjects were meeting a requirement of an introductory course in psychology. These subjects were tested in a different semester than those in Experiment I. The experimental and control subjects were assigned to conditions in groups of 4 as they entered the laboratory, on the basis of a predetermined, unsystematic order. The 4 basic lists at each frequency level were used once in all experimental and control conditions.

Results

–Serial recall–An analysis of variance for pair scores showed that recall increased reliably with list frequency [$F(2, 180) = 21.00$, $p < .01$], with the relevant pair-score means being .53, 1.72, and 2.08. Recall also increased reliably over the four list presentations from a pair mean of .75 to one of 2.14 [$F(1, 180) = 46.00, p < .01$]; and as would be expected, the average pair score of 1.70 for the forward-recall task was significantly higher than that of 1.19 for the backward-recall task [$F(1, 180) = 6.00$, $p < .05$]. However, there was no evidence in the form of interactions to suggest that backward- and forward-recall measures were not comparable in other respects.

The only other significant source of variation was the interaction of trial and frequency [$F(2, 180) = 5.00$, $p < .01$]. This interaction was based on the fact that recall for the lowest- to the highest-frequency lists following one list presentation was at means of .19, 1.09, and .97 respectively, while the corresponding means following four list presentations were .88, 2.34, and 3.19. This change over trials from a quadratic to linear trend is characteristic of prior findings (Turnage and McCullough, 1968), and the Trial 1 data replicate the curvilinear trend found in Experiment I. We now turn to an analysis of the association data.

–Free associations–An analysis of variance of the association data involving groups (controls; backward recall, one trial; backward recall, four trials; forward recall, one trial; and forward recall, four trials) as one main effect and frequency as another showed that the only significant source of variation was frequency [$F(2, 225) = 53.30$, $p < .01$]. The mean numbers of associations for the high-, medium-, and low-frequency lists respectively were 5.70, 5.20, and 3.57. Thus, there was no evidence in this analysis for modification of A–B associations as a function of learning conditions.

Since the free-association instructions directly encouraged the subjects to give both B and C associates to A, their associations were separated into these two classes. The mean number of C associates for the control subjects was .01, verifying that the lists were low in interitem associative strength at the beginning of serial learning, and, as expected, the mean number of C associates increased as a function of study trials for the experimental subjects. However, the minute size of this increase—from a mean of .02 to one of .10 over the four study trials—indicated that the subjects did not automatically transfer C associates into the free-association task. This impression was reinforced by a comparison of C means with pair-score means and by considering that on the average each subject had about 60 opportunities to give C associates.[1]

–Commonality of associations–In order to assess disruption of A–B by A–C associations, the following measure of commonality was employed. Each group of four subjects tested at the same time and under the same conditions was considered a block. Then, the number of different primary associations given by each block of subjects was divided by 48, which was the maximum number of different primaries possible for the 12 stimuli involved. Thus, the maximum range of commonality scores was from .25 (all subjects in the same block gave the same primaries) to 1.00 (no subject in the same block gave the same primary). If subjects in the same block failed to give any associate, which occurred quite rarely and only for low-frequency stimuli, this was scored as the 'same' response.

Table 2 shows mean commonality score as a function of groups and frequency. It can be seen there that, with one exception, stability of primary associations increased with frequency. This seems inconsistent with theoretical implications of the hypothesis of unit-sequence interference. That is, in principle at least, the most dynamic flux in serial order of associative output, and hence *in*stability of primary associations, ought to be for high-frequency stimuli.

An analysis of variance for the data in Table 2 showed that

Table 2. Mean commonality score as a function of groups and list frequency, Experiment II

	List frequency			
	Low	Medium	High	Row means
Control group	.88	.76	.76	.80
One presentation, backward recall group	.92	.85	.80	.86
One presentation, forward recall group	.97	.88	.82	.89
Four presentations, backward recall group	.96	.88	.77	.87
Four presentations, forward recall group	.93	.84	.75	.84
Column means	.93	.84	.78	.85

groups $[F(4,45) = 6.09, p < .01]$ and frequency $[F(2,45) = 49.78, p < .01]$ were significant sources of variance but that their interaction was not ($F < 1.00$). When Duncan's range test was applied to the means for groups (averaged across frequency levels), it turned out that the control condition differed from all of the experimental conditions at the .05 level. The only other means which differed in this analysis were those for the forward-recall groups having one versus four study trials prior to recall. In summary, then, the commonality analysis suggested that the serial-recall task had generally increased the variability of primary associates—which might reflect competition between A–B and A–C associations—but there was no evidence suggesting that such disruption varied with A's frequency in the language.[2]

DISCUSSION

Both Experiments I and II replicated a prior finding (Turnage and McCullough, 1968) of a quadratic function relating pair scores in short-term recall (Trial 1) and item frequency. Since this recall pattern is consistent with predictions from interference theory (see Underwood and Postman, 1960), it was expected that unit-sequence interference, in particular, would be in evidence as an important factor in disrupting serial recall of higher-frequency items. However, the present analyses of (*a*) changes in the quadratic trend relating recall and frequency as a function of conditions prior to recall (familiarization training, associative priming, or number of study trials) and (*b*) changes in the pool of associates to experimental stimuli as a function of learning and recall, indicated that associative interference was less 'powerful' than might be expected from theory.

Factors affecting quadratic trends

In Experiment I, it was expected that familiarization training prior to recall would 'straighten out' the quadratic trend relating frequency to recall, by reducing both letter-sequence and unit-sequence interference effects. The configuration of means in Table 1 support this prediction, and such an outcome seems consistent with interference theory. On the other hand, the fact that serial recall was facilitated by a free-association task involving the recall stimuli seems inconsistent with the theory. That is, if the inverse relationship between frequency and recall at the higher frequency levels involved competition between A–B and A–C associations, such competition should have been exaggerated by 'priming' A–B associations during the free-association task. However, the lack of an interaction of groups and frequency, and the configuration of means in Table 1, failed to support this prediction.

In Experiment II, changes in serial-recall patterns over trials were in accord with prior experimental data and at least compatible with interference theory. Yet the empirical evidence for the theoretical interplay between A–B and A–C associations was not particularly striking. That is, the form of the interaction of frequency and trials indicated that the quadratic function relating frequency and pair scores was 'straightened' over study trials, presumably due to the temporary inhibition or unlearning of competing A–B associations (see Underwood and Postman, 1960). Yet an immediate free-association test revealed few occurrences of A–C associations and, hence, little displacement of A–B associations by them. Of course, relative failures to give A–C associations could be interpreted as due to the persistent dominance of A–B associations, which is compatible with interference theory. However, such an explanation does not follow from the unlearning hypothesis being considered.

The commonality measures (see Table 2) also raise a pertinent question. That is, can the shifts in trend relating frequency and pair scores over study trials (which trend presumably mirrors attentuation of unit-sequence interference) be reconciled with the finding that increased variability of primary associations (which variability presumably involves A–B, A–C competition) was not dependent on A's frequency? An internally consistent theoretical account is certainly possible, but straightforward accounts on the basis of strict associative terms seem difficult to provide. Be that as it may, if the subject did 'unlearn' A–B associations while learning A–C associations, such modifications did not systematically vary with either A's frequency or the number of A–C study trials. Further, there was little evidence for automatic transfer of responses across the association and recall tasks for either Experiment I or II.

Response-pool integrity versus unit-sequence interference

What did seem automatic in both experiments was the subject's 'habitual' discrimination of response pools appropriate for association versus recall. Although the associative mechanisms by which such discriminations may be accomplished are difficult to specify, the present data require considering ways in which they might obviate unit-sequence interference. Underwood has, of course, expressed similar concern with respect to the so-called selector mechanism (Underwood and Schulz, 1960), with respect to the 'behavior' of laboratory versus linguistic associations (Underwood and Ekstrand, 1966), and with respect to attributes of memory such as 'between-test tags' (Underwood, 1969).

Similarly, Horton and Turnage (1970) found certain theoretical difficulties for interference theory in explaining empirically observed transfer (from a serial to a paired-associate task) without considering 'strategies' on the part of the subjects. Of particular relevance is their assumption that lists of serial items *may* be stored in memory in an ordered fashion without eliciting each other as a chain. In such cases, serial learning would be expected neither to produce a large number of A–C associations in an association task nor to disrupt A–B associations during A–C learning. Such an interpretation is certainly consistent with the present data, especially those from Experiment II. It also permits an interpretation of initial disruption of serial ordering for high-frequency lists (Experiments I and II) as due, in part, to the subjects' preliminary difficulty in discriminating the appropriate storage order for the greater multitude of high-frequency items generally available during recall (see also Turnage, 1970). In any case, the combined recall and association data indicate that a general theory of memory must be flexible enough to deal with other than automatic mechanisms of associative chaining if it is to remain viable (see also Underwood, 1969).

Notes

This research was partially supported by a grant to the Center for Language and Cognition from the Biomedical Science Support Committee of the University of Maryland, by a General Research Board grant from the University of Maryland to Thomas W. Turnage, and by Predoctoral Public Health Service Fellowship 1-F1-MH-37,686-01 from the National Institute of Mental Health to Jenny I. Steinmetz. Computer time for the project was supported by Grant NsG-398 from the National Aeronautics and Space Administration to the Computer Science Center at the University of Maryland. Requests for reprints should be sent to Thomas W. Turnage, Department of Psychology, University of Maryland, College Park, Maryland 20742. Received for publication May 4, 1970.

1. Data from an unpublished study by Thomas A. McCullough in 1969 showed only a *single* occurrence of a C associate out of more than 3,000 opportunities after up to four study trials followed by backward recall.

2. With respect to this general analysis, it may be noted that McCullough's unpublished study showed that the average disruption of primaries for high-frequency stimuli (*a*) was significantly less, following recall, for control (novel) stimuli than for experimental stimuli (.75 versus .81) and (*b*) *tended* to be less for control (no recall) than for experimental groups (.77 versus .80).

References

Horton, D. L., and Turnage, T. W. 1970. Serial to paired-associate learning: Utilization of serial information. *Journal of Experimental Psychology* 84:88–95.

Noble, C. E. 1952. An analysis of meaning. *Psychological Review* 59:421–430.

Postman, L. 1967. The effect of interitem associative strength on the acquisition and retention of serial lists. *Journal of Verbal Learning and Verbal Behavior* 6:721–728.

Slamecka, N. J. 1966. Differentiation versus unlearning of verbal associations. *Journal of Experimental Psychology* 71:822–828.

Thorndike, E. L., and Lorge, I. 1944. *The teacher's wordbook of 30,000 words.* New York: Teachers College, Columbia University.

Turnage, T. W. 1970. Proactive inhibition of initial items in short-term recall. *Psychonomic Science* 20:236–237.

Turnage, T. W., and McCullough, T. A. 1968. Letter-sequence and unit-sequence effects during learning and retention. *Journal of Experimental Psychology* 76:141–146.

Underwood, B. J. 1969. Attributes of memory. *Psychological Review* 76: 559–573.

Underwood, B. J., and Ekstrand, B. R. 1966. An analysis of some shortcomings in the interference theory of forgetting. *Psychological Review* 73: 540–549.

Underwood, B. J., and Postman, L. 1960. Extraexperimental sources of interference in forgetting. *Psychological Review* 67:73–95.

Underwood, B. J., and Schulz, R. W. 1960. *Meaningfulness and verbal learning*. Chicago: Lippincott.

RETROACTION, RECOVERY, AND STIMULUS MEANINGFULNESS IN THE A-B, A-Br PARADIGM

Edwin Martin and Stephen T. Carey
University of Michigan, Ann Arbor

The role of stimulus meaningfulness in paired-associate transfer, retroaction, and recovery was examined. Stimulus meaningfulness did not affect performance in second-list learning or in tests for retroaction in the A-B, A-Br paradigm. Relative to an A-B, C-B control paradigm, however, both transfer and retroaction were more negative with higher stimulus meaningfulness. Marginal absolute recovery was observed in the A-B, A-Br paradigm over repeated retroaction tests.

The experiment reported in this paper was planned and executed with the purpose of showing that spontaneous recovery is more marked with low-meaningfulness stimuli than with high-meaningfulness stimuli, as predicted by Martin (1968), and that spontaneous recovery is observable in the A-B, A-Br paradigm, as implicitly denied by Postman's list-differentiation theory of retroactive interference (Postman, Stark, and Fraser, 1968). In addition, we were intent upon further, confirming data on the role of stimulus meaningfulness in transfer and retroaction, a topic on which there is a paucity of information.

We employed a somewhat different test for retroaction and recovery than is ordinarily used, with an eye toward increased sensitivity. After learning trigram-digit pairs in either the A-B, A-Br or the A-B, C-B paradigm, the subject was administered a sequence of modified Brown-Peterson tests of short-term memory. The first task stimulus (an A) was presented, followed immediately by a rehearsal-preventing task that continued for either 0, 10, or 30 sec. Then the same stimulus was re-presented. The subject was to respond with the first digit that came to mind. The logic behind this procedure is that the initial, or activation, presentation of the stimulus should mobilize whatever underlying response-emission processes might obtain. The activation interval was varied so as to allow different amounts of time for these processes to produce a resolution of their effects.

The A-B, C-B paradigm was chosen as the proper control for the A-B, A-Br paradigm because they differ only in forward associations. The transfer factors of backward association, context association, response learning and availability, and such nonspecific effects as learning-to-learn and warm-up are identical in the two paradigms.

METHOD

-Materials–Two sets of nine CVCs (CUP, DEM, GYM, LAC, NAW, PEG, RON, SAW, YEL; BOM, FUL, JIM, KIS, MET, TAX, VAN, WIL, ZIP) and two sets of nine CCCs (BGN, DSK, GSP, JFN, KBH, PLS, RCK, TCH, XHF; CSF, FLN, GCZ, HMP, MBX, RBM, SGN, TPR, XLZ) were used as the high- and low-meaningfulness stimuli. Their mean association values were 96% and 52% respectively (Archer, 1960; Witmer, 1935). Each of these lists was paired in two different ways with the digits 1 through 9. The resulting four paired-associate lists for each level of meaningfulness were assigned to the A-B, A-Br and A-B, C-B paradigms equally often in all of the permissible permutations so that the meaningfulness levels of the stimuli were the same in both tasks and so that over subjects each list was the first-task and the second-task list equally often.

-Procedure–Learning in both tasks was by the aural anticipation procedure. All materials were recorded on magnetic tape (all stimuli were spelled) and presented to the subject over earphones. The prescribed response was presented 3 sec after the stimulus, the next stimulus was presented 3 sec after the preceding response, and the intertrial interval was 6 sec. First-list learning continued to a criterion of two perfect trials; second-list learning, to a single perfect trial. There was a new random ordering of the nine pairs on each trial. Those subjects in the A-B, A-Br paradigm were told, prior to A-Br learning, that the same stimuli and responses were to be re-paired; those in the A-B, C-B paradigm, that the stimuli would be new, the responses old.

During testing, each of the nine first-list stimuli was presented once in a modified Brown-Peterson short-term-memory situation. After the taped signal "Ready," a trigram was presented over the earphones. Three of these presentations were followed immediately by a soft *buzzer,* which in turn was followed by re-presentation of the same trigram, to which the subject was instructed to emit the first digit response that came to mind. This was the no-delay (2 sec) test condition, utilizing three of the nine trigram stimuli. For the other six trigrams, a *bell* sounded immediately after the trigram presentation, upon which signal the subject began work on a symbol-cancellation task. Then, either 10 or 30 sec later, the buzzer sounded, the trigram was re-presented, and the subject emitted a digit in response. He had 3 sec to respond with a digit before the next "Ready" was heard. The three delays

(2, 10, 30 sec) were distributed evenly, but unpredictably, over this sequence of nine tests.

After having tested 12 subjects in each combination of stimulus-meaningfulness level and paradigm, we decided to add 12 more subjects to each combination so that further tasks could be introduced. These latter 48 subjects were treated exactly as described above, except that additional, subsequent tests were also made. If the test procedure already described is called Cycle I, then Cycles II and III were exact replications except that over all three cycles each stimulus trigram was tested once at each interval. Thus for each combination of stimulus-meaningfulness level and paradigm, 24 subjects contributed to the Cycle I results, 12 to the Cycles II and III results.

The latter 48 subjects, upon completion of Cycle III, were then presented the first-list stimuli, one item at a time at a 5-sec rate, and asked specifically to respond to each with the response that went with it in first-list learning. This recall test was then followed by a similar test wherein the second-list stimuli were presented one at a time and the subject was asked to respond with only the second-list responses. For the A-B, A-Br condition, these two tests differed only in which responses were requested.

Subjects—The subjects were 96 undergraduates at the University of Michigan, women who volunteered for paid participation. They were assigned to the four experimental conditions randomly except for the restriction that at any given point in time the four groups had approximately equal numbers of subjects.

RESULTS

Learning

The mean (and median) number of trials to criterion in first- and second-list learning are shown in Table 1. Each entry summarizes 24 scores. As usual, first-list learning was more difficult with less meaningful stimuli. This difference was preserved in second-list learning in the A-B, C-B paradigm but was absent in the A-B, A-Br paradigm. Statistically, the effect of stimulus meaningfulness in first-list learning was only marginally significant $[t(94) = 1.75, .05 < p < .10]$. In second-list learning, it yielded $F(1, 92) = 5.82, p < .01$, and the interaction between meaningfulness and paradigm yielded $F(1, 92) = 6.75, p < .05$.

In the right-hand column of Table 1 are the means of individual-subject transfer scores[1] for the four conditions. The A-B, A-Br paradigm yielded negative transfer relative to the A-B, C-B paradigm only when the stimuli were of high meaningfulness. With low-meaningfulness stimuli, the nominal A-B, A-Br paradigm functionally approximated the A-B, C-B control paradigm. Analysis of variance on these individual-subject transfer scores yielded $F(1,$

Table 1. Mean (and median) trials to criterion and percent transfer

Stimulus meaningfulness	Paradigm	Trials to criterion List 1	Trials to criterion List 2	Percent transfer
High	C-B	8.4 (7.8)	4.6 (4.6)	42.5 (44.2)
	A-Br	8.2 (7.5)	5.9 (5.8)	22.1 (19.0)
Low	C-B	9.5 (9.7)	6.4 (6.2)	30.3 (38.0)
	A-Br	9.2 (9.0)	5.9 (5.7)	34.3 (34.5)

92) = 7.88, $p < .01$, for the interaction between stimulus meaningfulness and paradigm.

Test cycles

Shown in Figure 1 are the proportions for first-list responses emitted during Cycle I as a function of the length of the activation interval. Each point summarizes [(24 subjects) × (3 tests per interval)] 72 experimental events. This same pattern of results obtained in each of the three test cycles, except that the absolute heights of some of the functions increased over cycles.

If the A-B, C-B paradigm is taken as a control baseline, then clearly there was substantially more retroactive interference in the A-B, A-Br paradigm when the stimuli were of high meaningfulness than when they were of low. Collapsed over activation interval, the proportions of first-list responses out of 216 possible were .41 and .76 for the A-Br and C-B paradigms respectively with high-meaningfulness stimuli; the corresponding proportions with low-meaningfulness stimuli were .39 and .52. These figures give a 46% retroaction loss with high-meaningfulness stimuli, a 25% retroaction loss with low-meaningfulness stimuli. An analysis of variance for these Cycle I results yielded: for paradigm, $F(1, 92) = 37.56$, $p < .001$; for meaningfulness, $F(1, 92) = 11.35$, $p < .001$; and for their interaction, $F(1, 92) = 7.76$, $p < .01$.

With regard to possible spontaneous recovery over the 10- and 30-sec activation intervals, reference to Figure 1 shows there was absolutely none in the A-B, A-Br paradigm, to say nothing of differential amounts for low- and high-meaningfulness stimuli. The picture was irregular in the A-B, C-B paradigm. For these Cycle I data,

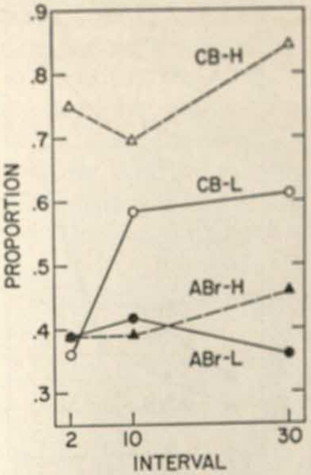


Fig. 1. Proportions of first-list responses during Cycle I as a function of activation interval for low-meaningfulness (L) and high-meaningfulness (H) stimuli

there was a main effect of interval [$F(2, 184) = 2.04$, $p < .05$]; but interval interacted neither with paradigm [$F(2, 184) = 1.26$, $p > .10$] nor with meaningfulness [$F(2, 184) = 1.42$, $p > .05$].

Consider now only those 48 subjects (12 per experimental condition) who participated in all three test cycles. In Cycles II and III, these subjects were retested with the stimuli presented in Cycle I (the first-list trigrams). We disregard activation interval here, since its effect in each cycle is as shown in Figure 1. Collapsed over intervals, the proportions of first-list responses emitted under each condition are shown in Figure 2 as a function of cycle number (ignore the curves as they extend to abscissa T). Each point summarizes [(12 subjects) × (3 intervals) × (3 tests per interval)] 108 experimental events.

The main effect of cycles was highly significant [$F(2, 88) = 14.92$, $p < .001$]; but the effect of cycles did not reliably interact either with paradigm [$F(2, 88) = 1.94$, $p > .10$] or with meaningfulness [$F(2, 88) = 1.60$, $p > .20$]. A separate analysis of just the A-B, A-Br paradigm indicated a marginal effect of cycles [$F(2, 44) = 2.38$, $.05 < p < .10$], no interaction between cycles and meaningfulness [$F(2, 44) = .70$], and clearly no effect of meaningfulness [$F(1, 22) = .26$].

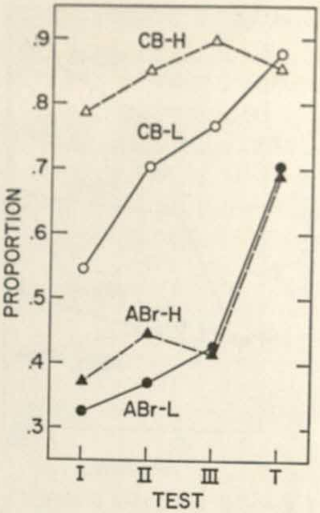


Fig. 2. Proportions of first-list responses as a function of test cycle for low-meaningfulness (L) and high-meaningfulness (H) stimuli

List-specific tests

At abscissa T in Figure 2 are shown the proportions of first-list responses emitted to their stimuli when the subjects were asked specifically for first-list responses. There are two properties of these data worth noting. First, for both low- and high-meaningfulness stimuli in the A-B, A-Br paradigm, emission of first-list responses increased 65% over the emission level observed in Cycle III. In Cycle III, the A-B, C-B paradigm was superior to the A-B, A-Br paradigm at the ratio of 2.0:1; in the list-specific test, the ratio was only 1.2:1. Second, in the A-B, C-B paradigm the effect of stimulus meaningfulness observed in Cycles I, II, and III was eliminated.

First- versus second-list responding

All data in this section are for the A-B, A-Br paradigm only. For Cycles I, II, and III respectively, the proportions of responses that were either a correct first- or a correct second-list response were .88, .88, and .92 with low-meaningfulness stimuli; for high-meaningfulness stimuli, .82, .85, and .86. Thus correct responding remained approximately constant over cycles, with a very slight overall superiority for low-meaningfulness stimuli (.89 versus .84).

The proportions of first- and second-list responses to low- and high-meaningfulness stimuli are shown in Figure 3. Over the three cycles there was clearly a competitionlike trade-off between the responses from the two lists. With low-meaningfulness stimuli, there appeared an easily discernible recession of second-list responses and concomitant recovery of first-list responses.

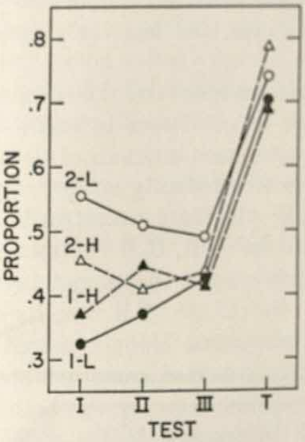


Fig. 3. Proportions of first-list (1) and second-list (2) responses as a function of test cycle for low-meaningfulness (L) and high-meaningfulness (H) stimuli in the A-B, A-Br paradigm

At abscissa T in Figure 3 are shown the proportions of responses emitted on the specific-list tests. The responses from both lists seemed more or less equally available, with a slight superiority for second-list responses. And the improvement in response emission on the final, list-specific test over that in Cycle III was commensurate for the two sources of responses, List 1 and List 2.

DISCUSSION

On the basis of this research, the following appear to be the effects of stimulus meaningfulness on paired-associate learning, transfer, and retroaction. (*a*) In first-list learning, performance is a direct function of stimulus meaningfulness.[2] (*b*) In second-list learning in an interference paradigm, performance is not related to stimulus meaningfulness (see also Martin, 1968). (*c*) Negative transfer in an interference paradigm relative to an appropriate control paradigm is inversely related to stimulus meaningfulness: the lower the meaningfulness, the more the interference paradigm behaves as the control paradigm (see also Martin, 1968). (*d*) In tests for retroaction in an interference paradigm, performance is not related to stimulus meaningfulness. (*e*) Amount of retroactive loss in an interference paradigm relative to an appropriate control paradigm is inversely related to stimulus meaningfulness: the lower the meaningfulness, the less the retroactive loss (see also Martin, 1968).

The picture for spontaneous recovery remains a study in frustration, at least insofar as one hopes to relate recovery to stimulus variables. Over the activation intervals of 10 and 30 sec in the present experiment, there was patently no recovery of any kind in the A-B, A-Br paradigm; yet there were trends that can only mean absolute recovery in the A-B, C-B control paradigm (see Figure 1). These latter trends suggest that something was indeed going on during the interval; but whatever it was, its effects netted to zero for the A-B, A-Br paradigm. Upon repeated testing in the A-B, A-Br paradigm, there occurred marginal absolute recovery (see Figure 2); but this seeming recovery was both less than in the A-B, C-B paradigm and, in contrast to the A-B, C-B paradigm, unresponsive to stimulus meaningfulness.

The most suggestive evidence of recovery is that shown in Figure 3 for the A-B, A-Br paradigm. Although the total number of responses available was approximately the same for low- and high-meaningfulness stimuli, in Cycle I there was for low-meaningfulness stimuli a clear dominance of second-list responses, a dominance that dissolved over cycles. In contrast, for the high-meaningfulness stimuli there was no particular list dominance and no change in list dominance over cycles. These effects seem to add up to the conclusion that low-meaningfulness stimuli are functionally less stable in an ambiguous test situation.

Observed performance on the list-specific tests (abscissa T in Figures 2 and 3) means to us that if the experimenter removes uncertainty about what he wants by way of response, the subjects can oblige him. In other words, even though substantial retroaction was observed during the test cycles, this retroaction is not traceable to broken or impaired stimulus-response associations.

We are prompted by these and other results to several tentative theoretic premises:

1. Unintegrated, fractionative nominal stimuli impose on the observer-experimenter an inherent uncertainty as to which of several possible ways his subjects will represent those stimuli in recognitive memory.[3] This uncertainty applies both to his observations of a single subject over re-presentations of the identical nominal stimulus and to those of a once-presented identical nominal stimulus over several or more subjects. For theoretical purposes, we are inclined to translate this observer-experimenter uncertainty into something called stimulus-encoding variability on the part of the subject.[4] From this premise follow the effects of numerous stimulus variables on the learning of a single paired-associate list.
2. During learning, the subject stabilizes his representation of the nominal stimulus. This means that part of learning is stimulus discrimination or elaboration, to the end that reliable stimulus identification is possible (see, for example, Ellis and Muller, 1964).
3. In a transfer task that involves identical nominal stimuli, the subject utilizes the results of stimulus-identification learning evolved in original learning, as opposed to reintroduction of coding variability (see Goggin and Martin, 1970). From this premise follows the result that although a stimulus variable is highly effective in first-list learning, it has little or no absolute effect in second-list learning (except as degree of first-list learning is pertinent).
4. If the stabilized, functional representation of the stimulus is multifeatured, then in transfer learning the subject may utilize only a subset of those features as retrieval cues for the second-list response, retaining the others as retrieval cues for the first-list response. This premise is loosely sufficient for explaining how it is that retroaction occurs in certain ambiguous test situations and how it is that on the specific-list tests of the present experiment the subjects could provide the required first-list responses. During the test cycles, the subject sampled the active features and then decided which response to emit on some basis of which we know very little (but see Postman, Stark, and Fraser, 1968; Martin and Mackay, 1970). On the specific-list tests, he sampled the active features and then de-

cided which response to emit on the basis of the experimenter's request.

This view involves a distinction not often made vis-á-vis the results of transfer and retroaction studies, namely, that although an experimental variable may affect *relative* transfer and retroaction, it may not affect transfer and retroaction-test performance per se. The observation that both transfer and retroaction are less negative with low-meaningfulness stimuli is an observation relative to a control condition; and it is clear in Table 1 and in Figures 1 and 2 that such a conclusion is due entirely to the fact that stimulus meaningfulness continues to exert its effects in the control conditions but not in the interference conditions.[5]

We proposed that in an interference paradigm, the subject uses in the second task the results of the stimulus-identification learning carried out in the first task. Given a reasonable criterion in first-list learning, one might suppose that this stimulus-identification learning has reached a point of stability; that is, that the functional codings of low- and high-meaningfulness stimuli are nearly equal in reliability. This supposition looks good in light of the absence of stimulus-meaningfulness effects in transfer and retroaction in the A-B, A-Br paradigm (see Table 1, Figures 1 and 2). But then why should there be a stimulus-meaningfulness effect in retroaction in the A-B, C-B control paradigm (see Figures 1 and 2)? And why should there appear a pattern of response dominance trade-off that is different for low- and high-meaningfulness stimuli in the A-B, A-Br paradigm (see Figure 3)?

We do not have ready answers to these questions. We pose them in order to express our opinion as to which of many theoretical problems are most important. This opinion is flavored by our conviction that the most neglected aspect of verbal-learning theory is investigation of stimuli as the point of origin of retrieval routes.

Notes

This research was supported by the Advanced Research Projects Agency, Department of Defense, and monitored by the Air Force Office of Scientific Research, under Contract No. AF-49(638)-1735 with the Human Performance Center, Department of Psychology, University of Michigan, Ann Arbor. Received for publication August 1, 1970.

1. $100(T_1 - T_2)/T_1$, where T_i is the number of trials to criterion on List 1 or 2.

2. This is a standard finding; see, for example, Underwood and Schulz (1960, p. 35 ff.).

3. For elaboration and discussion, see Höffding (1891), Rock (1962), and Martin (1968).

4. For direct experimental evidence, see Biederman, Dumas, and Lachman (1968), and Butler and Merikle (in press).

5. See Underwood and Ekstrand (1968). They attacked this problem using intralist stimulus similarity, with rather unrewarding, though suggestive, results.

References

Archer, E. J. 1960. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs* 74(10, whole no. 497).

Biederman, I., Dumas, J. S., and Lachman. 1968. The effects of stimulus uncertainty and S-R transitional probability on paired-associate learning. *Journal of Verbal Learning and Verbal Behavior* 7:864–868.

Butler, B. E., and Merikle, P. M. In press. Uncertainty and meaningfulness in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior.*

Ellis, H. C., and Muller, D. G. 1964. Transfer in perceptual learning following stimulus predifferentiation. *Journal of Experimental Psychology* 68:388–395.

Goggin, J., and Martin, E. 1970. Forced stimulus encoding and retroactive interference. *Journal of Experimental Psychology* 84:131–136.

Höffding, H. 1891. *Outlines of psychology,* trans. M. E. Lowndes. London: Macmillan.

Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review* 75:421–441.

Martin, E., and Mackay, S. A. 1970. A test of the list-differentiation hypothesis. *American Journal of Psychology* 83:311–321.

Postman, L., Stark, K., and Fraser, J. 1968. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior* 7:672–694.

Rock, I. 1962. A neglected aspect of the problem of recall: The Höffding function. In *Theories of the mind,* ed. J. Scher, pp. 645–659. New York: Free Press of Glencoe.

Underwood, B. J., and Ekstrand, B. R. 1968. Differentiation among stimuli as a factor in transfer performance. *Journal of Verbal Learning and Verbal Behavior* 7:172–175.

Underwood, B. J., and Schulz, R. W. 1960. *Meaningfulness and verbal behavior.* Chicago: Lippincott.

Witmer, L. R. 1935. The association value of three-place consonant syllables. *Journal of Genetic Psychology* 47:337–360.

EFFECT OF LEVEL OF MANIFEST ANXIETY AND TYPE OF PRETRAINING ON TIMING BEHAVIOR

Devendra Singh
University of Texas at Austin

High- and low-anxiety subjects were trained on a timing schedule of differential reinforcement of low rate of responding, either without pretraining or after facilitatory or interfering pretraining. The low-anxiety subjects performed better than the high-anxiety subjects in all three conditions.

A great many studies have shown that subjects who differ in level of anxiety, as measured by the Manifest Anxiety Scale (Taylor, 1953), perform significantly differently on eyelid-conditioning and verbal-learning tests. There has, however, been no study investigating the effect of level of anxiety on complex operant conditioning. If it could be demonstrated that low anxiety and high anxiety have differential effects on operant conditioning, it would add generality to the Manifest Anxiety Scale as a selective device.

To this end, high-anxiety and low-anxiety subjects were trained on a timing task with a schedule of differential reinforcement of low rate of responding (DRL). Furthermore, since animal studies show that performance on such a schedule is affected by the nature of pretraining (Schmaltz and Isaacson, 1966), the role of pretraining was also investigated. Briefly, the experimental design consisted of giving high- and low-anxiety subjects either interfering pretraining (fixed-ratio schedule, which generates a high rate of responding), facilitatory pretraining (fixed-interval schedule, which generates a lower rate of responding with 'scallops'), or no pretraining.

METHOD

-Subjects–The subjects were selected from an introductory psychology course on the basis of their scores on the Manifest Anxiety Scale. The students within the upper 25% of the score distribution were designated high-anxiety subjects, and those in the lower 25% of the distribution were designated low-anxiety subjects. A total of 68 subjects (34 from each anxiety group) was selected from 300 students taking the test. All subjects were randomly assigned to one of the following experimental conditions: no pretraining (14 subjects from each group), fixed-ratio pretraining (10 subjects from each group), and fixed-interval pretraining (10 subjects from each group) prior to DRL-schedule training. The ratios and intervals were set at 15 and 15 sec.

-Apparatus–A 20-sec DRL with a 5-sec limited hold was programmed. A multiple-stimulus projector (Industrial Electronics) was mounted at eye level about 5 ft from the subject's chair. A telegraph key, wired through programming devices, was mounted on the arm of the chair. If the subject pressed the key within the 20–25-sec interval, a plus sign was flashed on the projector for .3 sec; if he pressed the key earlier than 20–25 sec or did not press within the interval, a minus sign was flashed for .3 sec. For fixed-ratio pretraining, a plus sign was flashed at the end of the ratio; for fixed-interval pretraining, all premature presses produced a minus sign. All plus and minus signs were superimposed on a dim green background light on the projector screen. The green background light on the projector screen was activated at the start of the experiment and remained on throughout the experimental session. The number of correct and incorrect responses and the interresponse times were recorded. The experimental room was illuminated by a 100-w light bulb that was painted black (except 1 sq in.) and mounted 9 ft directly above the projector. Two exhaust fans in the room produced a masking noise.

-Procedure–Upon arrival, the subject was taken to the experimental room by a male experimenter. Instructions were recorded on a tape recorder. All the subjects were told that the experiment was a time-judgment task and that every time they guessed the right time interval (by pressing the key) a plus sign would flash on the screen and two cents would be added to a bank of fifty cents; a wrong guess would be indicated by a minus sign and two cents would be subtracted. The subjects were also told that they could use any strategy to figure out the time interval except counting. At the end of the instructions, the subjects in the no-pretraining condition were trained on the timing task (DRL-hold schedule) for 60 min. The other subjects were permitted to obtain 25 reinforcements on either the fixed-interval or the fixed-ratio schedule, depending on the experimental condition to which they belonged. At the end of the pretraining, the subject was informed that henceforth a new time interval would be used but that the rest of the task would remain the same. Then he was trained on the timing task (DRL-hold schedule) for 60 min. Verbal reports were obtained from each subject at the end of the experimental session.

RESULTS

Performance

The efficiency (reinforcement/response) ratios for successive 5-min blocks were calculated, for each subject, to determine progressive performance on the timing task. Three separate analyses of variance for repeated measures, one for each pretraining condition, were performed on arc-sine transformed $(X + 2)$ reinforcement/response ratios.[1] Figure 1 presents the mean arc-sine reinforcement/response ratios along with significant probability level for differences (based on a two-tailed t test) for the three experimental (pretraining) conditions for successive 5-min blocks.

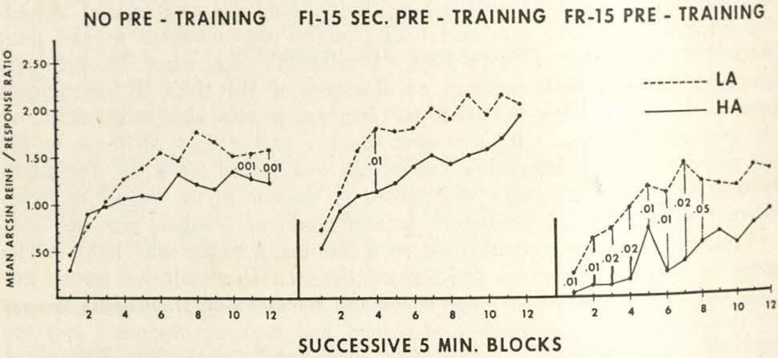


Fig. 1. Performance on the timing task for low-anxiety (LA) and high-anxiety (HA) subjects in the three experimental conditions

As Figure 1 shows, the performance of low-anxiety subjects was better than that of high-anxiety subjects in all three experimental conditions. Under the no-pretraining condition, there were no performance differences between the anxiety groups during the first 50 min; after that, however, the performance of the low-anxiety subjects became significantly better than that of the high-anxiety subjects. Compared to the no-pretraining condition, fixed-interval pretraining improved the performance of both anxiety groups; but in spite of this facilitatory pretraining, high-anxiety subjects did not perform better than low-anxiety subjects. As a matter of fact, the performance of the low-anxiety groups in the fixed-interval condition was somewhat, though not significantly, better than that of the high-anxiety groups.

The most dramatic performance differences became evident in the fixed-ratio condition. First, this condition produced poor performance by both anxiety groups. For example, the terminal performance of both low- and high-anxiety subjects after fixed-ratio pretraining was significantly poorer than with no pretraining [for low-anxiety subjects, $t = 7.89$, $df = 22$, $p < .001$; for high-anxiety subjects, $t = 10.98$, $df = 22$, $p < .001$] and significantly poorer than with fixed-interval pretraining [for low-anxiety subjects, $t = 8.73$, $df = 22$, $p < .01$; for high-anxiety subjects, $t = 6.74$, $df = 22$, $p < .01$]. Second, the low-anxiety subjects in the fixed-ratio condition started out performing better than the high-anxiety subjects in the same condition, and they maintained significantly superior performance for 40 min. Thereafter, they continued to perform better than the high-anxiety subjects, but the differences were not significant.

Interresponse time

The interresponse times were recorded for the intervals 0–5, 5–10, 15–20, and 25 (overshoot) for successive 5-min blocks for each subject. Again because of heterogeneity of variance, all scores were transformed to log scores before performing repeated measures of variance. The results indicated significant differences between the low- and high-anxiety subjects only for the 0–5-sec time interval; thus, Figure 2 presents the mean log responses under the 0–5-sec interval along with probability levels for differences (based on a two-tailed *t* test) for the three experimental conditions.

As is evident from Figure 2, under the condition of no pretraining both low- and high-anxiety subjects responded approximately equally throughout the experimental session. Though there was a reduction in the mean number of responses as the session progressed, there was no significant trend difference between the high- and low-anxiety subjects. Even when a facilitatory fixed-interval pretraining was given, both anxiety groups responded right after reinforcement. Compared to high-anxiety subjects, the low-anxiety subjects made fewer errors, though the differences were not significant. However, after fixed-ratio pretraining, both anxiety groups responded right after reinforcement compared to no pretraining as well as to fixed-interval pretraining. After fixed-ratio pretraining, the low-anxiety subjects responded significantly less

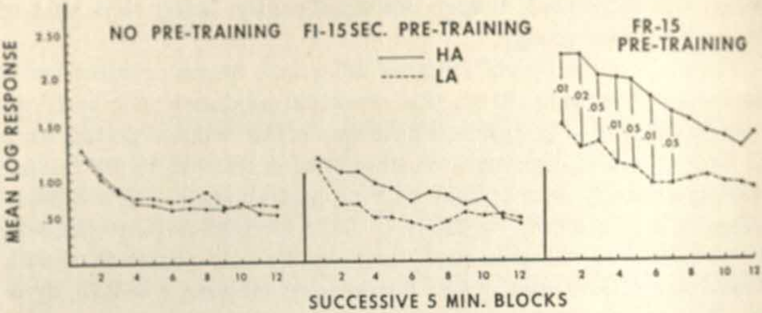


Fig. 2. Interresponse time in the 0–5-sec interval for low-anxiety (LA) and high-anxiety (HA) subjects in the three experimental conditions

than the high-anxiety subjects for 35 min; thereafter, the differences were not significant. Thus, it appears that the poor performance of the high-anxiety subjects after fixed-ratio pretraining was due to greater responding in earlier time intervals. As training continued, the high-anxiety subjects suppressed premature responses; though within the experimental session, they never performed better than low-anxiety subjects.

DISCUSSION

The main findings of the present experiment are that both level of anxiety and pretraining can affect performance on a timing task with a schedule of differential reinforcement of low rate of responding. If anxiety has drive properties, as suggested by Spence (1958), it is interesting that—unlike the results of animal studies—the level of drive did here affect performance on the task. The performance of low-anxiety subjects was superior to that of high-anxiety subjects in all three pretraining conditions, though the performance differences became strikingly large after inhibitory pretraining. One of the reasons that no significant effect of facilitatory pretraining was observed could be the limitation imposed by its fixed-interval schedule. It can be argued that this pretraining really produced an interfering tendency, since the time interval used in the timing task was different than that used in pretraining. Possibly, facilitatory effects of pretraining might have been observed at a fixed interval of 20 sec.

Although the complexity of the present experimental arrangement places it beyond the boundaries of the theoretical analysis of Taylor (1953) and Spence (1958), the superiority of low-anxiety subjects to high-anxiety subjects in all three conditions appears to contradict the theory. Nonetheless, the present findings provide additional support that subjects selected on the basis of widely different scores on the Manifest Anxiety Scale differ in their performance on complex as well as simple tasks—in the present case, a complex task involving timing behavior.

Notes

This research was conducted at North Dakota State University and was supported by research funds of the Graduate School. The author is grateful to David Rehovsky and Terry Robinson for their help in conducting the experiment and to Professor Abram Amsel for his critical reading of the manuscript. Received for publication April 28, 1970.

1. The transformation of scores was necessary because of lack of normalities under all three conditions.

References

Schmaltz, L. W., and Isaacson, R. 1966. The effects of preliminary training conditions upon DRL performance in the hippocampectomized rat. *Physiology and Behavior* 1:175–182.

Spence, K. W. 1958 A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *American Psychologist* 13:131–141.

Taylor, J. A. 1953. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology* 48:285–290.

BOOK REVIEWS

Julian Hochberg
Columbia University

Memory and Attention

By Donald A. Norman. New York: Wiley, 1969. Pp. ix, 201. $8.95; paperback, $4.95.

In this book an 'information processing' approach is adopted in which the human is treated as a complex system actively trying to impose organization on its sensory input. On such a view the most important (and most interesting) psychological processes are selective attention, pattern recognition, memory storage, and retrieval from memory. The reader is presented with a clear statement of contemporary positions on each of these topics, and the exposition is frequently illuminated by excerpts from the original sources.

The book has a distinctly modern flavor. This is perhaps due to the strong emphasis on efforts to model psychological processes. The models discussed in some detail range from Broadbent's filter theory of attention, through an electronic model of word association. Although the chief emphasis of this book is clearly on current theory, Norman has shown a serious regard for historical antecedents. This is most clearly exemplified in the long discussion of the ancient art of mnemonics—a topic that is all too easily sloughed off as 'gimmicky.'

The book is appropriate for a course on almost any level, since little psychological background is assumed. It could serve as an excellent supplementary text (especially in the paperback edition) for courses in perception, cognition, memory, or learning. Although the book's length precludes complete coverage, Norman makes a special point of providing the reader with useful suggestions for further reading.

Howard Egeth, *Johns Hopkins University*

Models of Thinking

By F. H. George. London: Allen & Unwin, 1970. Pp. 166.

This book on artificial intelligence asks how one might program a digital computer to model human concept formation, hypothesis formation, logical deduction, and language in one continuous process. It selects from the computer models available in the 1960s some components for a composite program, sketching out the assumptions underlying such work and describing the programs by use of flow diagrams.

The book succeeds in its attempt to present in a clear way a selection of computer models of thinking as information processing, but the author himself points to the weakness of the book: "Unwittingly we may present rigorous but oversimplified models and theories, and as a result the work will either be ignored as inept or rendered much less useful than one might hope" (p. 15). However, to readers sympathetic to the computer-model approach to thinking, the book may be a useful summary of much of the work done in the past decade.

Herbert F. Crovitz, *Durham Veterans Administration Hospital*

Thinking as a Science

By Henry Hazlitt. Rev. ed. Los Angeles: Nash Publishing, 1969. Pp. iii, 162. $4.95.

This book, originally published in 1916, is reissued with a new epilogue that lists the changes the author would make were he to begin afresh, and with minor editorial changes that tend to diminish the force of the original. Hazlitt wrote this book when he was 21 years old, before he succeeded to H. L. Mencken's editorship of the *American Mercury* and before his long career as a journalist specializing in economics.

Hazlitt writes in the epilogue that he would prefer now to call the book *The Art of Thinking*. Indeed this would be more appropriate, for there is no science in it, no mention of research studies on thinking, nothing but 'wisdom.' It is full of merely plausible recommendations, unsupported by empirical research, on matters such as the disciplining of attention, on making one's reading efficient and energizing, and on getting one's ideas into words and down on paper. In his chapter on concentration he writes: "If the reader is one of those rare and fortunate beings whose thoughts run only to useful subjects, and who always concentrate from pure spontaneous interest, I sincerely advise him not to force himself. And if such a being happens to be reading the present chapter I assure him he is criminally wasting his time and that he should [drop the book or] turn to the next chapter with all possible haste" (p. 50). The material in brackets appeared in the 1916 edition but was deleted from the 1969 edition. It may yet be good advice for the readers of this journal.

What ought one to demand of a book on the art of thinking? Perhaps an exploration of "how far the knowledge accumulated by modern psychology can be made useful for the improvement of the thought-processes of the working thinker." This was the goal that Graham Wallas stated for himself in the preface to his great book on *The Art of Thought* (1926). Wallas wrote with a flair for high comedy, and like Hazlitt he was a man of affairs. He *used* Hazlitt's masterful comparison of the intellectual methods of Herbert Spencer and John Stuart Mill to explore "that particular 'stroke' which will bring the energy of the organism most easily and most competently to bear on his task." After noting that William Morris wore himself out making confused

Marxist speeches on street corners, Wallas reported: "Bernard Shaw said to me, 'Morris has come into this movement with all of his energy, but not with all of his intellect'" (Wallas, p. 156).

A similar brush paints both Hazlitt and modern psychology. We have been lax in bringing our growing technical knowledge of thinking to bear on how best to go about the task of thinking for practical accomplishment or even for maximal enjoyment. Hazlitt writes in his epilogue that were he to begin again he would lean on the work of George Polya (*How To Solve It*, 1954), and he does mention the work of Graham Wallas, although the name is spelled "Wallace." It seems likely that we can do better; if we do, the least that can be said of this reissue of Hazlitt's book is that it reminds us that the job has not yet been done in a satisfactory way.

Herbert F. Crovitz, *Durham Veterans Administration Hospital*

The Origins of Intellect: Piaget's Theory

By John L. Phillips, Jr. San Francisco: W. H. Freeman, 1969. Pp. xviii, 149. $2.75.

Three Theories of Child Development

By Henry W. Maier. 2d ed. New York: Harper & Row, 1969. Pp. 342. $8.25.

Piaget and Knowledge

By Hans G. Furth. Englewood Cliffs, N. J.: Prentice-Hall, 1969. Pp. 270. $7.95.

Studies in Cognitive Development: Essays in Honor of Jean Piaget

Edited by David Elkind and John H. Flavell. New York: Oxford University Press, 1969. Pp. 503. $10.00; paperback, $4.50.

These four books represent the diversity available to the reader who desires an interpretation of Piaget. During the early phases of the dramatic upsurge of interest in Piaget's work, there were at most one or two secondary sources to which the uninitiated American reader could turn for an overview and interpretation. That is no longer the case, and the reader can now critically and selectively choose his 'guide to Piaget.' To this reviewer, the most basic and critical dimension for that choice is the extent to which the interpretation is both founded in and explicitly addresses itself to the theory of development that Piaget proposed. While the need for such characteristics would seem to go without saying, they are often found to be conspicuously absent.

Piaget's theory of the development of cognition can be characterized at the most general level in terms of two related propositions. The first is the idea that cognition or knowing is organized in terms of a structured set of internalized mental *acts*, or operations, the organization of which becomes transformed in development. The second specifies the unifying functional mechanisms by which this reorganization occurs. Generated from this theoretical framework have been Piaget's admittedly fascinating characterizations of children's thought. In the absence of a consideration of the mental operations *producing* these modes of thought, however, the products themselves are of considerably less interest or import. Both Phillips's and Maier's treatments of Piaget are representative of this unfortunate tendency in the numerous interpretive summaries of Piaget now available: they present a descriptive typology of stages of childhood thought without adequately considering, in particular, the mental operations producing a given concept or, in general, the theoretical mechanisms Piaget proposed to underlie the progression of stages. Phillips presents in outline style a brief definition of the major concepts in the Piagetian system, but he does not present the theoretical system itself with sufficient depth or coherence to allow an appreciation either of it or of the long sequence of stages he then proceeds to outline.

Maier's summary of the theoretical system is longer but likewise fails to give the reader a coherent picture. The six generalizations that, according to Maier, summarize Piaget's concept of development (p. 102) convey a picture neither of the organized structure of mental actions nor of the mechanisms in terms of which this structure becomes transformed. Moreover, several of Maier's generalizations are in fact notably inaccurate. Both play and language, he suggests, are vehicles for adaptation; whereas Piaget has clearly specified play as pure assimilation and language as a symbolic medium for the exercising of operations. Most misleading, he claims that in Piaget's views "experience rather than maturation defines the essence of cognitive development" (p. 94); whereas the concept of the organism-environment interaction as the basic source of patterning is clearly at the heart of Piaget's theory of the mechanisms of development. Maier's characterization of Erikson's developmental theory and of learning theory according to Robert Sears, while not notably inaccurate, suffers from the same difficulties as his characterization of Piaget. The essence of the theories as theories of the nature of development is not communicated with sufficient coherence to make meaningful the description of the various stages, sequences, or outcomes that the theories specify.

Both Phillips and Maier attempt to apply the theories they review, Phillips to education and Maier to psychotherapy. Because the essential elements of the theories as theories of development have not been well clarified, however, their applications remain superficial. Maier, for example, ends up suggesting that the three theories—organismic, psychoanalytic, and behavioristic—are readily integrated into a more total picture of the developing organism, a picture which can then be of use in clinical evaluation. But because he has failed first to differentiate the theories in terms of basic issues pertaining to the nature of development, his integration of the theories is a superficial one.

Furth's approach is a refreshing one. He presents Piaget as a philosopher asking a number of philosophical questions: "What is the nature of knowledge?" "How is knowledge constructed?" "What is the relation between the subject and object of knowledge?" In considering these questions and Piaget's solutions to them, Furth manages to convey a comprehensive and coherent picture of Piaget's system and of the fundamental theoretical issues it addresses. Most significant, he skillfully presents Piaget's theory in its context of both biology and epistemology. The relation between biological and cognitive adaptation, he makes clear, is more than one of analogy; cognitive functions emerge as a specialized subsystem of the general biological system of organism-environment interchanges or regulations. Also skillfully handled are Furth's discussions of equilibration and the theoretical relation between development and learning, the relation between figurative and operative knowing, and the constructive character of symbolic representation. The technique of interspersing the exposition with brief readings from Piaget's own writing is an effective one. The excerpts from *Memoire et intelligence* and *Biologie et connaissance* should be especially interesting to American readers, as these works are not yet available in English. The only fault that might be found with Furth's presentation is that he to some extent does the opposite of what Phillips, Maier, and most writers on Piaget do: he presents a picture of Piaget as a theoretician reflecting on a number of theoretical questions and fails to give the reader an adequate sense of all of the empirical knowledge that Piaget's investigations have produced. What the reader loses is not just the rich and fascinating description of Piaget's stages of childhood thought. More important, he gains no sense of the delicate interplay between empirical finding and theoretical construction that has characterized Piaget's work nor, more generally, of the contributions that empirical science and epistemology can and must make to each other, a matter which Piaget himself has stressed.

Elkind and Flavell's collection of essays may mark the advent of what promises to be a new stage in American psychologists' study of cognitive development. While the quality of the essays is uneven and the collection as a whole suffers from a lack of unity, several of the articles, due to the kinds of problems with which they concern themselves, represent a significant advance. Specifically, the two major theoretical contributions in the collection, one by Pinard and Laurendeau, and the other by Flavell and Wohlwill, suggest by the nature of their presentations that we who are interested in cognitive development, and particularly in Piaget's theoretical approach to the subject, can now go beyond merely attempting to 'interpret' Piaget correctly—that we can go on to consider in their own right some of the theoretical issues he raised. These kinds of discussions are a refreshing advance over the 'interpretations' to which most writers on the Piagetian approach to cognitive development have limited themselves. Pinard and Laurendeau, for example, present an interesting discussion of the difficult problems of horizontal and vertical "decalage," how the two concepts relate to one another, and how these developmental "asynchronisms" mesh with the "structured whole" concept of development. Flavell and Wohlwill's article, though somewhat less clarifying, discusses the characteristics of stage transition in terms of the underlying structures and of the behavioral or empirical indicants; most important, it considers the strategic problem of how structural transformations can be inferred from behavioral ones. Two other articles are of particular interest as good summaries of Piaget's general viewpoint and current research in two areas: Inhelder summarizes the Genevan work on memory, and Sinclair the work on language development and on language and thought. Other articles in the collection are of lesser interest and, particularly in the case of those on the conservation problem, vary in the extent to which they are founded in an understanding and appreciation of the general Piagetian view of development.

Deanna Kuhn, *Columbia University*

An Analysis of Personality Theories

By Albert Mehrabian. Englewood Cliffs, N. J.: Prentice-Hall, 1968. Pp. 217. $7.95.

As the title of this book indicates, it is a meta-theoretical work. Professor Mehrabian elaborates a system whose purpose is to allow the reader to analyze any given personality theory. The main contention is that the underlying metaphor or metaphors of any given theory must be grasped in order to understand it. He argues that these underlying metaphors largely determine the observations the theorist will make and heavily influence the nature of the conceptualizations he will utilize to explain the interrelationships among the phenomena he observes.

The use of metaphor is defined as the application of a familiar and known system to a new and unknown one. Freudians are said to have applied the metaphor of the hydraulic system, a known and familiar system, to the understanding of personality dynamics, a relatively new and unknown area. The cognitive theorist Piaget is described as having applied the biological metaphor of adaptation to the field of cognitive development. The metaphor of biological adaptation (a description of the intake of nutrients and the elimination of waste materials through assimilation and accommodation) is used as an analogy to describe the process of cognitive development. Metaphor, then, seems to mean analogy, and the contention is that the analogies that are used heavily constrain the nature of personality theories. Professor Mehrabian proceeds to analyze Rogerian, psychoanalytic, instinct-need-habit-trait-factor, and cognitive-developmental approaches in terms of their underlying metaphors.

There are other important features to this book. Of particular interest is a chapter on computer-simulation approaches to the study of personality. It is pointed out that computer simulation is not a theory but a method of testing and elaborating any theory. The author also prescribes ways of judging the relative merits of theories and suggests categories for the description of personality-related phenomena.

Although some outstanding undergraduates may well profit from this book, it is clearly most suitable for graduate students. Familiarity with the theories discussed must be derived from other sources, as suggested by the author. The book itself, however, can be comprehended without much knowledge of the theories discussed. This is a useful work that can profitably be used in conjunction with other materials in a course on personality theory. It would also be one excellent source of information for graduate seminars that touch on the nature of theory in the field of personality and related areas.

Robert D. Singer, *University of California, Los Angeles*

ANNOUNCEMENTS

The Committee on International Exchange of Persons announces that applications for senior Fulbright-Hays awards for lecturing and research during 1972–1973 in over 75 foreign countries will be accepted in the spring of 1971. Specialists in the field of psychology who are U.S. citizens and have a doctorate or college teaching experience are invited to indicate their interest in 1972–1973 awards by immediately completing a simple registration form, available on request from the Senior Fulbright-Hays Program, 2101 Constitution Avenue, Washington, D. C. 20418. Registrants will receive the detailed announcement of available awards as soon as it is issued in the spring, in time to consider the possibilities and to apply before the closing date.

Senior Fulbright-Hays awards generally consist of a maintenance allowance in local currency to cover normal living costs of the grantee and family while in residence abroad, and round-trip travel for the grantee (transportation is not provided for dependents). For lecturers going to most non-European countries, the award includes a dollar supplement, subject to the availability of funds, or carries a stipend in dollars and foreign currency, the amount depending on the assignment, the lecturer's qualifications, salary, and other factors.

a good choice for your personality and adjustment course

PSYCHOLOGY OF PERSONALITY AND ADJUSTMENT

BY FRANZ A. FREDENBURGH

A complete treatment of both classical and modern personality theorists. Covers all major theorists through Maslow and examines their theories in light of current research. Extensive photographs and illustrations; chapter organizers. Publication date, March 1971 $9.50 (tent.) Approx. 650 pages

a comprehensive introductory reader

ENCOUNTER

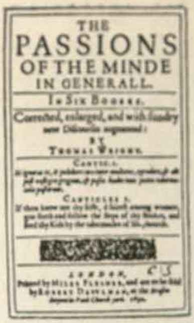
Issues of Human Concern

BY ROBERT V. GUTHRIE

Ideally suited to an introductory course or courses dealing with personality and adjustment. Contains selections relevent to today's student, showing the impact of psychological theory as it is in the "real world". Includes an introduction to each section and selection. Accompanied by a test booklet. 27 articles and case studies. $3.95. 306 pages

Please write for a complimentary copy in exchange for course information.

CUMMINGS PUBLISHING COMPANY, INC.
2727 Sand Hill Road
Menlo Park, California 94025



THE
PASSIONS
OF THE MINDE
IN GENERALL.
In Six Bookes.
Corrected, enlarged, and with sundry new Discourses augmented:
BY
THOMAS WRIGHT.

THE PASSIONS OF THE MINDE IN GENERALL

THOMAS WRIGHT

A Reprint Based on the 1604 Edition with an Introduction by Thomas O. Sloan

Wright's "pithie discourse about the passions of the minde" was one of the first originally English works on the physiological "humours" and the passions corresponding to them. Published in five editions between 1601 and 1630, his book has long been regarded as an important work in the history of psychology. Wright believed that studying the passions should contribute to man's efforts to lead a life of Christian virtue, but unlike other philosophers of his age, he included within his study of the passions lengthy discussions of rhetoric.

Sloan's introduction provides a brief biography of Wright and points out the importance of *The Passions* as a clear reflection of the cross-currents of political, religious, psychological, and rhetorical movements of its time. $12.95.

BIOLOGICAL AND SOCIAL FACTORS IN PSYCHOLINGUISTICS

EDITED BY JOHN MORTON

These nine essays discuss the extent of the uniqueness of human language and speculate on its origins. They present valuable new data from comparative study of the brain, the anatomy, and the behavior of living primates and of man.

The book is the result of a symposium held as part of the XIXth International Congress of Psychology in London during the summer of 1969. The contents are: "The Roots of Language" by Bernard Campbell, "Can Humans Talk?" by John C. Marshall, "Explaining Linguistic Universals" by David McNeill, "Comparing and Contrasting" by Roger Wales, "What Could Possibly Be Innate?" by John Morton, "The Grammar of Sign Language and the Problems of Language Universals" by I. M. Schlesinger, "Social and Natural in Semiotics" by A. A. Leontiev, "Syntax and Semantics" by Raqaiya Hasan, and "The Integrated Study of Language Behaviour" by Thomas G. Bever. $8.95.

UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Revised Edition

Samuel A. Kirk, James J. McCarthy, and Winifred D. Kirk

The Revised ITPA is an individually administered diagnostic test which taps significant abilities and disabilities in psycholinguistic functions in children between the ages of two and ten. It evaluates abilities in three dimensions:

☐ channels of communication (auditory-vocal and visual-motor)
☐ psycholinguistic processes (receptive, organizing, and expressive processes)
☐ levels of organization (automatic and representational)

Norms are presented for each of 12 subtests and in scaled scores. In addition, the test as a whole can yield: (1) a Composite Psycholinguistic Age and a Psycholinguistic Quotient, and (2) an estimated Binet mental age and an IQ. The complete kit includes directions for administration and scoring, 2 picture books of test materials, newly designed chips, tray, and pictured sequences for the Visual Sequential Memory Test, dispensable picture strips and scoring tissues for the new Visual Closure Test, 6 objects for the Verbal and Manual Expression Tests, a 33⅓-rpm instruction record, and 25 record forms. $43.50.

A FILMED DEMONSTRATION OF THE ITPA

Produced by Winifred D. Kirk and Samuel A. Kirk

Sixteen mm. black-and-white sound film (43 minutes). $360.00

THE DEVELOPMENT AND PSYCHOMETRIC CHARACTERISTICS OF THE REVISED ITPA

John N. Paraskevopoulos and Samuel A. Kirk

Describes the model for and rationale of the tests. Paper, $5.50.

UNIVERSITY OF ILLINOIS PRESS Urbana · Chicago · London

DISEASES OF ATTENTION AND PERCEPTION

Monte Jay Meldman, M.D.
Director of Research
Forest Hospital
Des Plaines, Illinois

241 pages integrating data on attention and perception in behavioral pathology into a classification system of diseases of attention. 580 references review the important literature.

Available from your local book store and from
Pergamon Press, Inc.
Maxwell House
Fairview Park
Elmsford, New York 10523

Price: $13.50

---

from Cambridge

EXPLANATION IN THE BEHAVIOURAL SCIENCES

A Confrontation of Viewpoints

Edited by R. BORGER and F. CIOFFI

Philosophers, psychologists, sociologists and linguists often differ sharply in their assessment of the types of explanation characteristic of and appropriate to the behavioural sciences. With this in mind, the editors invited a number of distinguished philosophers and behavioural scientists to debate a series of connected problems. Each chapter takes the form of an exposition, followed by a comment and then by the expositor's reply. $15.00

CAMBRIDGE UNIVERSITY PRESS
32 East 57th Street
New York, N. Y. 10022

media sociology: a reader

Edited and with an introduction by Jeremy Tunstall

In "The Structure of Foreign News," in this volume, Johan Galtung and Mari Holmboe Ruge point out that international action is based on the image of international reality, and ". . . the regularity, ubiquity, and perseverance of news media . . . make them first-rate competitors for the number-one position as international image-former." Therefore, "research into the adequacy of the image the news media give of the world is of primary importance."

Intended for students of social science, this book adds a valuable comparative aspect to media studies. It focuses on the British system of communication and British audiences (though nearly half the articles are non-British in subject matter or authorship). The twenty-five essays concentrate less upon the audience than most media studies and more upon the communications organizations and communicators, such as journalists and TV producers.

One-third of the articles have not been previously published. The volume includes: "Approaches to the Historical Development of Mass Media Studies" by Roger L. Brown; "Britain's National Media Pattern" by Winston Fletcher; "The Impact of TV on the Audience for National Newspapers, 1945-68" by James Curran; "Public Service and the Private World" by Tom Burns; "Developments in Soviet Radio and Television News Reporting" by Gayle Durham Hollander; "The Mass Media and Egyptian Village Life" by Ibrahim Abu-Lughod; "Books, Book Readers and Bookshops" by Peter H. Mann; and "The Audience for Election Television" by Jay G. Blumler and Denis McQuail.

Other contributors are Joseph Bensman, Oliver Boyd-Barrett, Ronald F. Bunn, David Butler, Melvin L. DeFleur, Philip Elliott, B. P. Emmett, James Halloran, D. Harper, Hilde T. Himmelweit, Michael Lane, Aleksander Matejko, Joan Munro, Robert Silvey, Donald Stokes, Andrew Tudor, and Malcolm Warner. $12.50.

UNIVERSITY OF ILLINOIS PRESS Urbana · Chicago · London

from COLUMBIA UNIVERSITY PRESS

Stroke A Slain Warrior

FRANK MICHAEL CORTINA. From tape recorded therapy sessions with addicts, the author has selected 17 which form the individual chapters of this book. Each chapter tells the story of one person and why and how he or she became addicted. The vignettes, fascinating in themselves, were chosen to illustrate a wide diversity of types involved in drug addiction. $6.95

Mental Retardation and Its Social Dimensions

MARGARET ADAMS. The author emphasizes that the social worker, once equipped with the necessary information about mental retardation, can make a contribution through the various disciplines of social work to the best adjustment of the retarded individual, his family, and all those with whom he comes in contact. She sets forth and illustrates with case material the use of social work techniques in dealing with these problems.

Introduction by Joseph M. Reid, Director of the Child Welfare League of America. Foreword by Stanley Davies.

"Studies of the Child Welfare League of America" $10.00

Address for orders: 136 So. Broadway, Irvington, N. Y. 10533
In Canada: McGill-Queen's University Press, 3458 Redpath St.,
Montreal 109

"What forces us to believe that our only choice is between war or surrender? What is there about human thinking under stress that pushes every group conflict toward mutual annihilation? . . . Let us look into the dynamics of our own minds."

AN ALTERNATIVE TO WAR OR SURRENDER

by Charles E. Osgood

Here is a brilliantly conceived, carefully blueprinted plan for an approach to international relations in the nuclear age that is as fresh now as it was when the book was first printed in 1962 because it is psychologically sound. It offers a way out of the dangerous dilemma posed by the increasing tensions of the arms race that is leading to a permanently divided world, if not a fatal nuclear war.

This plan, which the author has named *Graduated Reciprocation in Tension-reduction* — GRIT — assumes that the problems we face are primarily matters of human nature and human relations and that the Soviet people and their leaders are more like us than the "bogey man" our psycho-logic creates.

It provides that the United States take the initiative in reversing the arms race and in reducing world tensions through a carefully planned and executed "peace offensive" designed to invite reciprocation from the Communist bloc. It is a step-by-step approach with built-in safeguards aimed at maintaining our security.

"Let us look into the dynamics of our own minds, because only through understanding ourselves can we ever hope to understand others." IB-7. 183 pages. Paper, $1.45.

UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

27

Vol. 84, No. 2 Founded in 1887 by G. Stanley Hall June 1971

Akc
Docket & issue
to the librarian
R Das
16/11/71

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys **M. E. Bitterman**
Charles W. Eriksen **Benton J. Underwood**

With the cooperation of Karl M. Dallenbach, William Epstein, Rand Evans, Robert Fox, Wendell Garner, J. P. Guilford, Harry Helson, E. R. Hilgard, Julian Hochberg, Francis W. Irwin, F. Nowell Jones, John Krauskopf, G. L. Kreezer, W. C. H. Prentice, Allan Wagner

FRED ATTNEAVE and RICHARD K. OLSON. Pitch as a Medium: A New Approach to Psychophysical Scaling 147

MUKUL K. DEY. Distributed Practice in Paired-Associate Learning: Interstimulus Similarity 167

KENT GUMMERMAN. Selective Perception and the Number of Alternatives 173

GLORIA J. FISCHER. Sum of Event Probabilities and Subjective Estimates of π in Probability Learning 180

DENNIS J. DELPRATO. Specific-Pair Interference on Recall and Associative-Matching Retention Tests 185

continued on back cover

The American Journal of Psychology was founded in the interest of general experimental psychology and devotes its pages to reports of original research, to shorter notes and discussions, and to book reviews pertinent to the area. Manuscripts may be submitted to any one of the following:

LLOYD G. HUMPHREYS
Psychology Building
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820

M. E. BITTERMAN
Department of Psychology
University of South Florida
Tampa, Florida 33620

CHARLES W. ERIKSEN
Psychology Building
University of Illinois at Urbana-Champaign
Champaign, Illinois 61820

BENTON J. UNDERWOOD
Department of Psychology
Northwestern University
Evanston, Illinois 60201

All books intended for review should be sent to the editor of that section:

JULIAN HOCHBERG
Department of Psychology
Columbia University
New York, New York 10027

Contributors are asked to follow the style described in the *Publication Manual of the American Psychological Association.* They should keep abstract and title quite brief, use footnotes only when essential, avoid all but the most commonly recognized or proper abbreviations (2 ft, 4 deg, ANOVA, BD-IX rats; *but* subjects, observers, experimenter, paired-associate learning, afterimage), and use the tables and figures to supplement rather than repeat the text. Authors of articles, notes, and discussions are given 100 complimentary copies without covers; covers and additional copies may be ordered at the time galley is returned.

The subscription price is $10.00 a year (foreign $10.50); single copies cost $2.75. Back numbers and volumes may also be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which address all inquiries of a business nature (subscriptions, advertisements, changes of address) should be directed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

© 1971 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

-3 JAN 1972

Vol. 84, No. 2 *continued from front cover* June 1971

WAYNE LEE. The Effects of Expected Value Difference and Expected Regret Ratio on Preference Strength 194

STEFAN SLAK. Univariate and Bivariate Information Storage 205

JOSEPH GLICK. Is Autokinetic Motion Related to Experienced Displacement? 210

LAUREN HARRIS and M. JOSEPH SCHALLER. Form and Its Orientation: Reexamination of a Child's-Eye View 218

ROBERT PASNAK. Pattern Complexity and Response to Distinctive Features 235

MARTIN BOLT. Imagery in Associative Learning and Memory 246

RALPH Y. SASSON. Semantic Organizations and Memory for Related Sentences 253

LARRY E. WOOD and JAMES V. HINRICHS. Cuing of Recall Order in Immediate Memory with Different Rates of Presentation 268

G. J. JOHNSON. Transfer as a Function of Paradigm and Interspersed-Pair Difficulty 276

Book Reviews 288

Books Received 299

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS
Urbana, Illinois 61801

PITCH AS A MEDIUM:
A NEW APPROACH TO PSYCHOPHYSICAL SCALING

Fred Attneave and Richard K. Olson
University of Oregon

Certain psychophysical continua have a primarily morphophoric function: they are the media in which patterns occur. The appropriate scale for any such medium is one that reflects the invariances preserved in transposition behavior. The results of two experiments designed to scale pitch by transposition methods lead to the conclusion that the musical scale, not the mel scale, describes the morphophoric function of pitch. However, an abrupt breaking point is evident at about 5,000 Hz, above which transposition behavior is erratic.

Psychophysics has a curiously autonomous status in contemporary psychology. The measurement of elementary sensations is obviously important if sensations are viewed as the building blocks of perception. Since this view has not been widely entertained for some five decades, however, one is entitled to wonder why psychophysicists go on measuring sensations with undiminished zeal. The basic fact that perception is relationally determined—that perceptual objects owe their identity to certain relational invariants of patterns or configurations—has profound consequences for the psychophysical treatment of such commonly investigated continua as brightness, loudness, and pitch. A major defect of conventional psychophysics is its failure to take these consequences into account.

Consider the example of pitch, with which this article is specifically concerned. Rather diverse methods (summated *jnd*s, interval judgments, ratio judgments) yield fairly consistent results for pitch: essentially, the *mel* scale. (See Stevens, Volkmann, and Newman, 1937; Stevens and Volkmann, 1940.) What these methods have in common, for all their diversity, is that they treat individual tones as perceptual objects, or psychological units, each of which is conceived to have its characteristic pitch sensation, or value on the *attribute* of pitch. Things are different outside the laboratory: people rarely treat individual tones as auditory units or tie particular behaviors to particular pitches, nor would it be adaptive for them to operate in such a manner. A very simple case of an auditory unit with realistic object properties is that of a brief sequence of tones, as in a melodic phrase. It is well known that such a sequence is *transposable:* the pattern as a whole may be raised or lowered in pitch without destroying its perceptual identity. The corollary of this statement is that pitch is a *medium,* in which the same pattern may have different locations.

An appropriate scale of pitch, so considered, is one that describes its morphophoric, or form-bearing, properties. The simplest such scale would be one on which any two perceptually equivalent patterns differed only by an additive constant. Let us inquire whether a scale of this kind exists for pitch, and if so, what psychophysical transformation on frequency will yield it.

If musical conventions are to be taken at face value, a logarithmic transformation gives the desired scale. Adding a constant to the log frequency of every note in a melody may change its key, but not its musical identity. However, the musical scale is not in agreement with the mel scale; a musical interval near 6,000 Hz contains more than six times as many mels as the 'same' interval near 100 Hz. This disagreement might be interpreted in a variety of ways. One possibility is that the musical scale is right, and the mel scale wrong, with respect to pitch as a medium. At the other extreme, it might be argued that musical equivalence is more or less arbitrary—being based perhaps on the structure of musical instruments, perhaps on relationships between harmonics of tones sounded together—and that under some ideal conditions people would actually transpose on the mel scale. A third possibility, of course, is that the scale optimally describing equivalences among patterns remains undiscovered.

In the experiments reported here, our purpose was to infer a morphophoric scale of pitch from transposition behavior obtained under conditions as free of irrelevant influences as possible. To this end we employed patterns consisting of pure tones, presented successively, at a level (40 db) low enough to render aural harmonics negligible over a major portion of the range investigated.

EXPERIMENT I

The general plan of the first experiment was to present the subject with tonal patterns from various parts of the frequency spectrum, and have him reproduce these patterns in a standard middle region. Each pattern presented was made up of only two frequencies, no more than an octave apart, alternated with variable duration to form some distinctive, repetitive temporal configuration.

Method

—Subjects—Six paid student volunteers from the University of Oregon were selected to provide different levels of musical experience and ability. We were particularly interested in whether the subject had any effective verbal structure for the classification of intervals in musical terms. The two subjects we designate as *musical* were able to identify common musical intervals (either by names like 'third,' 'fifth,' and so on, or by assigning the component notes to a sol-fa scale), and likewise to construct intervals with pure tones, given their values. One of these, JF, was a female voice major who played the piano and who had sung leading roles in several local musicals. The other, DM, was a male graduate student in psychology who sang and played the guitar. None of the four subjects designated *nonmusical* had any definite verbal structure for describing intervals. One of them, however, RB, a male graduate student in psychology, had studied piano briefly as a small child, was strongly interested in music, and spent a great deal of time listening to both 'classical' music and jazz. The remaining three, KN, SD, and LC, were female undergraduates none of whom had more than a very casual interest in music. Three other nonmusical subjects were dropped from the experiment because they failed to meet a minimum performance criterion.

—Materials and apparatus—The stimuli are described in essentially musical terms. Each tone is represented by a pair of numbers such as 2-5 or 6-12. The first number designates the octave in which the tone occurs; the second its position within that octave on a 12-tone or chromatic scale, each step of which is exactly one-twelfth of a $\log_2$ frequency unit. Middle C (261.6 Hz) is called 3-1; thus 3-10 is the A above middle C (440 Hz), 4-10 is the A an octave higher (880 Hz), and so on. The tones employed covered a seven-octave range, from 1-7 (92.5 Hz) to 8-7 (11,840 Hz).

The 'stimulus' to be matched on a given trial was always a temporal pattern of two such tones within an octave of each other. The sequencing mechanism consisted of a phonograph turntable (rotating at 33 rpm) on which rested a cardboard disc extending beyond its edge. This disc was divided into 16 equal sectors, any combination of which could be cut away beyond the edge of the turntable to provide a photoelectric cam. Transparent sectors allowed a beam of light to strike a photocell; opaque sectors interrupted it. The photocell in turn controlled a circuit that switched back and forth between the two tones without temporal overlap, but without a noticeable interval between the alternating tones. Switching clicks were not altogether eliminated by the circuitry but were held to an unobtrusive level.

Typical patterns (in terms of duration of successive tones) were such cycles as 113146113146..., 222226222226..., and so on. Each of the two tones occurred three times in every cycle; there were six alternations per cycle. However, the tones were not necessarily equated for total duration; therefore every trial on which a pair of tones was pulsed by a particular cam disc was counterbalanced by another trial on which the same pair was pulsed by the 'negative' of that disc, so that the higher and lower tones exchanged places. In all, 24 discs were used, 12 basic patterns and their 'negatives.' Assignment of tone pairs to discs was random, except for the counterbalancing noted above. The aim in using these rhythmic cycles was to make the object perceived a pattern or configuration, even though only two frequencies were involved. (That this aim was not altogether unrealistic is illustrated by such a familiar example as the first four notes of Beethoven's Fifth Symphony.)

Given a standard or 'stimulus' pattern of this kind, the subject's task was to match it (to construct a pattern as much like it as possible) in a different frequency region. He did this by adjusting a single oscillator, since one tone of the comparison or 'response' pattern was fixed by the experimenter and served as an anchor. By means of a hand switch, the subject could freely choose at any time during a trial to listen to either the stimulus pattern or the response pattern. This switching operation was on the inputs to the photoelectric cam mechanism; therefore the stimulus and response patterns were identical in their temporal or rhythmic aspects.

Over trials, 84 different pairs of tones were employed in stimulus patterns: 12 intervals (1–12 chromatic steps) were presented in each of seven octaves. The midpoint of the interval was always a C (2-1, 3-1, . . . , 8-1) in the case of even-numbered intervals, or half a chromatic step higher in the case of odd-numbered intervals. The response pattern was invariably anchored in the 5-1 to 6-1 octave, the region between 1,046.5 and 2,093 Hz. The distance between the anchor tone in this pattern and its counterpart in the stimulus pattern was always an integral number of octaves minus half an octave (six chromatic steps). This means that a strictly musical transposition demanded the greatest possible change in key, as from C to F-sharp.

The tones making up the stimulus patterns were generated by a Grayson-Stadler Twin 950C oscillator. Those in the response patterns were produced by two RCA WA-44C oscillators. One of the latter was located in front of the subject in a sound-attenuated room; all other major items of equipment were located in a separate control room with the experimenter. No attempt was made to calibrate the oscillators accurately for frequency. Instead, all frequencies were determined by a vernier stroboscope, Conn Instrument's Strobocon 6T4. This remarkable device provided a degree of precision far beyond the requirements of the present study, even if one liberally discounts the manufacturer's claim of ±.05% accuracy.

It seemed advisable to present all tones at a constant loudness. The sensation level of 40 db was chosen as a compromise between two objectives: to hold aural harmonics to an insignificant level, and to make the tones loud enough to be heard comfortably. (Results were also spot-checked at 20 db.) To maintain constant loudness, we needed to know what voltages to apply to the output tranducers—a pair of Koss Pro-4 headphones—at the various frequencies represented in stimuli. This information was obtained directly (that is, without separately considering the transfer functions of phone and ear) by having four preliminary subjects adjust in amplitude each of the 85 frequencies employed (on two trials, one ascending, one descending) to match the loudness of a 1,000-Hz tone 40 db above threshold. The four resulting curves were in excellent agreement with one another, and were averaged to provide the values of output voltage subsequently used. Amplitudes of the two stimulus tones and the anchor tone of the response pattern were controlled manually by the experimenter. Uniform loudness of tones from the subject's oscillator was achieved by a mechanical linkage of controls: a spring-loaded arm from the amplitude knob rode on a carefully machined bronze cam attached to the frequency dial. A cardboard disc, also attached to the frequency dial, hid most of this mechanism as well as the frequency markings on the dial.

—Procedure—Subjects were first checked for adequate sensitivity over the experimental frequency range. They were asked if the highest tone used seemed about as loud as the 1,000-Hz reference tone at 40 db. They were then spot-checked throughout the experimental range to ensure that they had no lacunae. Their ability or inability to identify musical intervals was also ascertained, as described. The task was then presented to them as simply a matching of 'patterns' or 'tunes.' They were never told that they should match patterns by musical standards, or that they should refrain from doing so, but merely that whatever sounded like the same pattern was the desired response.

Subjects were familiarized with the task in a series of practice trials. The first 12 practice patterns, given on days 1 and 2, were drawn from the two octaves overlapping the response octave and extending half an octave beyond it in either direction. Feedback was given with special care to avoid imposing any particular scale on the subject's judgments. Thus, he was asked to reconsider his response if it was grossly beyond the limits of a mel match, a musical match, or anything in between; this was quite practical in view of the small range. Also, he was specifically shown his error if he reversed the ordinal relationship of the tones or if he set his tone to the same frequency as the corresponding stimulus tone. Twelve more practice trials, on days 3 and 4, were drawn randomly from the whole experimental range. On these, the subject was corrected only for qualitative errors: ordinal reversals and identity or octave matches. Three subjects who were unable to eliminate such qualitative errors (and moreover showed exceedingly high variability) were dropped after the practice trials.

In the body of the experiment, the subject was given a total of eight trials (two series of four) on each of the 84 stimulus pairs. The four trials within a series represented the possible combinations of two procedural variables: (*a*) whether the subject adjusted the high or the low tone of the pair, and (*b*) whether he began the trial with his oscillator set at the high or the low end of its range. These (4 × 84) trials were given in a random order, with the restriction that a given combination of the two variables just mentioned not occur twice in succession. Assignment of temporal patterns (photoelectric cams) to trials was also random. The whole series was then repeated in a new random order, but the photoelectric cam used on any given trial was the 'negative' of that used on its counterpart in the first series.

The subject was told to listen first to the stimulus pattern and then adjust the response pattern to match it, switching back to the stimulus pattern whenever necessary. Each experimental session lasted about an hour and included 20 trials, more or less, depending on the subject's speed and endurance. He was allowed as much time as he needed to make a setting.

Results and discussion

The eight responses a subject made to each stimulus pair were averaged, thus yielding a mean transposed interval value. (The averaging was all done on a log frequency scale because of the relative uniformity of variability in these terms. The means so obtained are not systematically different from medians; see Experiment II.) These values are plotted for each subject in the six graphs of Figure 1. The abscissa of each graph is essentially a log frequency scale, and a given 'column' represents stimulus intervals that have the same logarithmic midpoint.[1] The subject's matching interval (in the 5-1 to 6-1 response octave, 1,046.5–2,093 Hz) is shown on the ordinate in terms of chromatic steps, or twelfths of an octave, with a separate function for each stimulus interval labeled in similar terms. A graph representing perfect transposition by musical standards would thus consist of 12 horizontal straight lines intersecting the ordinate at the correspondingly numbered points.

The two musical subjects, JF and DM, showed very similar behavior. To a good approximation, they transposed on a logarithmic or musical scale over the first five octaves. In the two highest octaves their curves, as shown in Figure 1, remain fairly flat for small intervals (up to about five chromatic steps), but for larger intervals a gross breakdown occurs. In the sixth (next to highest) octave, JF markedly compresses the three largest intervals and DM the seven largest ones; this dip is so extreme that in several cases a large interval is reproduced smaller than one only half its size. In the highest octave, DM shows a surprising recovery, or decompression of the large intervals (though his ordering is still by no means perfect); JF shows a similar decompression of intervals 10, 11, and 12, but a new compression of intervals 6, 7, and 9.

Both JF and DM reported that high tones (above about 5,000 Hz) lacked the *musical quality* characteristic of lower ones. Considering that the higher tone of a pair increased in frequency as interval size increased about a given midpoint, we can inquire

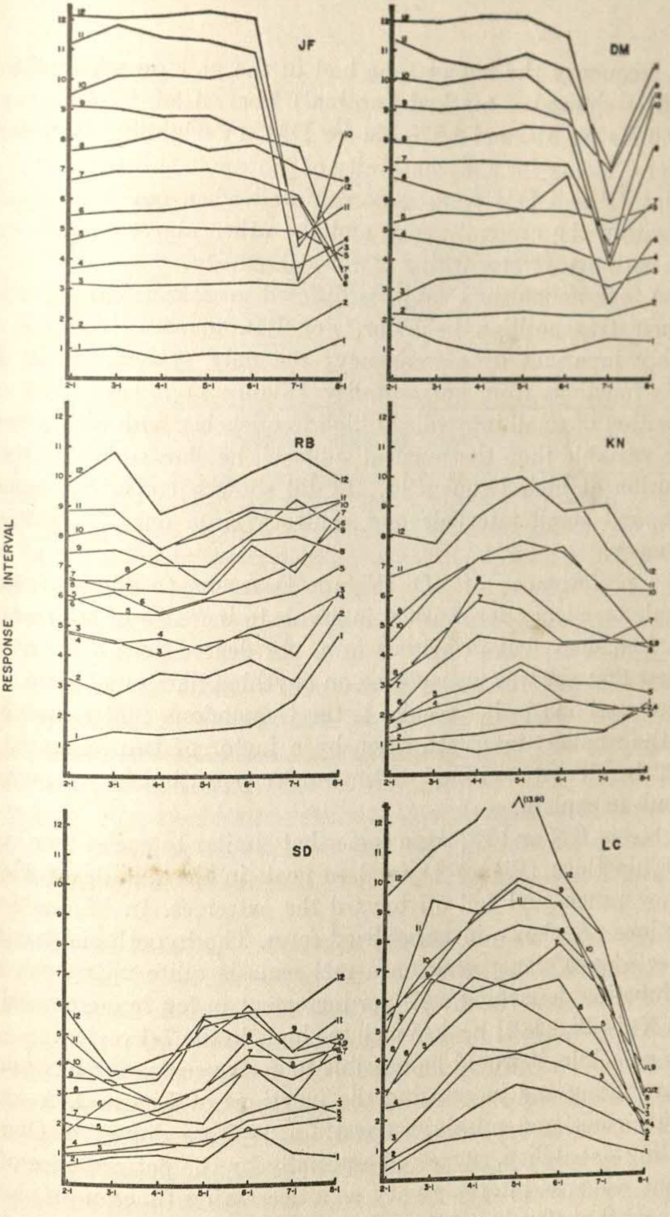


Fig. 1. Transposition functions for individual subjects in Experiment I. Each curve shows matching responses to a particular stimulus interval of 1 to 12 chromatic steps presented at various frequency levels. Points on the abscissa are an octave apart and represent the midpoints of stimulus intervals; 3-1 is middle C, 261.6 Hz.

what frequency the higher tone had in the pair on which the subject first showed a marked (ordinal) aberration. This value was 5,587.6 Hz for JF, and 4,978 Hz for DM. (Variability data provide further evidence for a discontinuity of process in this region.) It may be that JF and DM were most confused when one tone of a pair was within the musical range and the other above it, and less so when both tones were lacking in musical quality.

The four nonmusical subjects differed greatly from one another in their transposition behavior. For RB, musical intervals were roughly invariant over frequency; the only systematic deviation of his functions from horizontality (Figure 1) reflects some overestimation of small intervals at high frequencies. Although generally more variable than the musical subjects, he showed relatively less confusion at high frequencies. He did show a consistent tendency to expand small intervals and contract large ones, regardless of frequency.

The performance of SD (Figure 1) seems to show a modest, irregular tendency for musical intervals to increase in apparent size with frequency, but in neither form nor degree does this variation suggest that she was transposing on anything like a mel scale. What is most striking in her results is the tremendous contraction of all but the smallest intervals, often by a factor of two or more. This might be viewed as a kind of time-order error, but its magnitude is difficult to explain.

Subjects KN and LC were somewhat similar to each other in that their functions (Figure 1) tend to peak in the middle of the frequency range, and fall off toward the extremes. In Figure 2 their functions are shown in normalized form. The hypothesis that these subjects were transposing on a mel scale is quite untenable, since the function describing mels per increment in log frequency (shown with Xs in Figure 2) peaks much too high, in the 7-1 region.

In principle, it would be possible to derive scales for KN and LC by averaging and integrating the functions of Figure 2. A variety of suspicions and misgivings restrain us from doing so. One disquieting possibility, suggested especially by the performance of LC, is that response intervals vary as a decreasing function of the distance of the stimulus pattern from the response region. If this were the only difficulty, it could be remedied by a counterbalancing of stimulus and response regions. Of more fundamental concern, however, are the indications that SD, KN, and LC were employing idiosyncratic judgmental or response heuristics inconsistent with

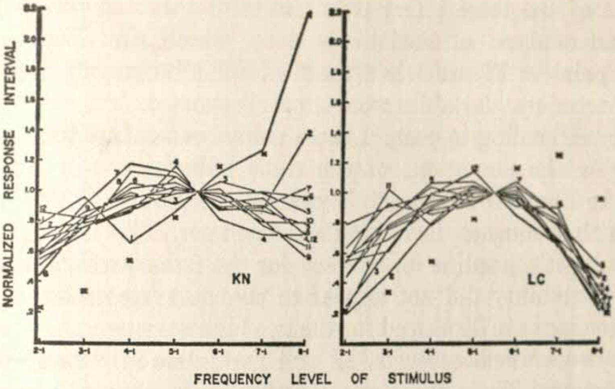


Fig. 2. Normalized functions for subjects KN and LC, Experiment I. Each curve from Figure 1 has been divided by the mean of its values at 5-1 and 6-1, so that all curves intersect in the middle of the response region. Musical transposition would be represented by a horizontal through 1.0; mel-scale transposition by the unconnected Xs.

the task assigned them, which was to make the response pattern sound as much like the stimulus pattern as possible. Note, for example, that a scale based on SD's intervals 2, 3, and 4 (Figure 1) would have little resemblance to one derived from her intervals 10, 11, and 12: the former functions rise steadily with frequency, more than doubling between 2-1 and 7-1, whereas the latter seem to follow a peculiar sinusoidal course, ending up about where they begin.

Again, note that KN's transposition of interval 1 is markedly different from her transposition of all the others (Figures 1 and 2); also that her functions for intervals 6, 7, and 8 peak at a lower frequency than those for intervals 9, 10, 11, and 12. Any such interaction with interval size, whatever its source, clearly violates the assumptions of the scaling procedure we set out to employ. These and other considerations led us to the alternative paradigm described in Experiment II. Meanwhile, let us not lose sight of the fact that the musical subjects JF and DM gave us a close approximation to musical transposition, except in the two highest octaves, and that the behavior of RB (who, it may be recalled, was a devoted music listener, though he had no verbal categories for intervals) was also quite reasonably consistent with the musical scale.

Some of the most interesting results of this experiment emerge from an analysis of variability data, which are now described. Every point in Figure 1 is a mean of eight judgments, over which three procedural variables were counterbalanced. A variability measure corresponding to each of these points was calculated simply by averaging the deviations of the eight judgments from their own mean. In Figure 3 the variability scores are plotted against interval size, with frequency level as the parameter. The heavy curve in each subject's graph is an average for the five lowest octaves, over which variability did not appear to change systematically;[2] separate functions are displayed for the two highest octaves.

The two musical subjects, JF and DM, showed remarkably similar patterns. Their functions for the lower octaves are approximately flat; that is, variability did not increase with interval size. Over this region, variable error of frequency settings is of the order of 2% or less, but in the two highest octaves it becomes tremendously greater. Note that both subjects showed an abrupt increase in variability, by about a factor of five, at exactly the same place in the sixth octave—at the transition from interval 6 to interval 7. This is the step in which the high tone of the pair increased from 4,978 to 5,274 Hz. (Interval size was increased by alternately raising the high tone and lowering the low tone; at this particular transition the low tone remained unchanged at 3,520 Hz.) Thus we have further evidence for a discontinuity of process at just about 5,000 Hz.

We have observed already that JF and DM showed no increase in variability with interval size over the lower five octaves: for these functions the correlations are at a pure chance level, .18 and −.05 respectively. In the top octave, however, the correlation increases to .34 for JF, which is suggestive though not significant, and to .74 ($p < .01$) for DM.

The situation is quite different for the nonmusical subjects, all four of whom showed definite increases in variability with interval size in the lower octaves, the correlations being .68 ($p < .02$) for RB, .76 ($p < .01$) for SD, .88 ($p < .001$) for KN, and .94 ($p < .001$) for LC. In none of these subjects do we observe the dramatic increase in variability at high frequencies that characterized the musical subjects, though SD does appear somewhat more variable in the top octave, and the quasi-musical RB further shows some weak suggestion of discontinuity in the 5,000-Hz region. Note that at high frequencies nonmusical subjects are actually less variable

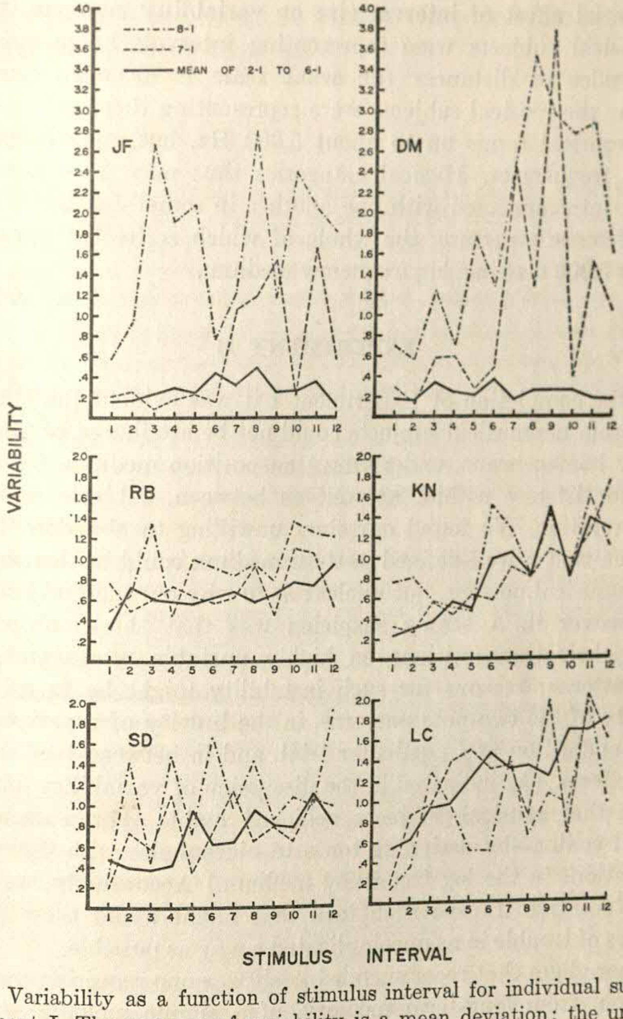


Fig. 3. Variability as a function of stimulus interval for individual subjects, Experiment I. The measure of variability is a mean deviation; the unit is a chromatic step, which represents a constant *proportional* change (about 6%) in frequency. Interval magnitude is likewise shown in chromatic steps on the abscissa. The five lowest frequency levels are averaged (heavy line); the two higher octaves are plotted separately.

than musical subjects, who were perhaps trying vainly to apply musical standards to nonmusical tones.

These data strongly support the view that intervals may be represented or encoded in either an analog or a digital manner. The differential effect of interval size on variability suggests that the nonmusical subjects were representing intervals as quantities or magnitudes or distances (on what scale is an open question), whereas the musical subjects were representing them in qualitative or categorical terms up to about 5,000 Hz, but quantitatively at higher frequencies. Musical categories thus may be conceived as rigidly interconnected with one another in something like a solfeggio reference structure, the whole of which is freely transposable (below 5,000 Hz) in a log frequency medium.

EXPERIMENT II

At the completion of Experiment I it was evident that those results from nonmusical subjects could not be attributed to properties of any homogeneous, underlying transposition medium, for reasons of inconsistency within, as well as between, subjects under this interpretation. We found ourselves unwilling to abandon the supposition that a well-defined pattern medium could be demonstrated in nonmusical persons, but unclear as to why our method had failed to discover it. A strong suspicion was that those subjects had based their transpositions on highly unstable representations of the patterns; reasons for such instability might be found in the poverty of the two-note patterns, in the brevity of experience with the standard on any particular trial, and in between-trial interference effects. (As indicated in the discussion of variability data, it is likely that musical subjects achieved stable representations by digital coding—by assigning tones to pigeonholes with determinate separations in the log frequency medium.) Accordingly, we sought an alternative transposition task that might avoid these possible sources of trouble in as uncomplicated a way as possible.

A paradigm that recommended itself was one requiring the transposition, from long-term memory, of a simple melodic sequence, already highly overlearned, in which the component notes had always had the same frequency values during learning. Such a sequence is not easy to think of, since a familiar tune may have been heard in many keys, but it occurred to us after some consideration that the NBC chimes might meet our requirements quite well. Upon checking with the National Broadcasting Company, we were assured that the chimes had never varied in key; they are always composed of the notes G, E, and C, ending on middle C. We proceeded then with an experiment in which the subjects were asked to reconstruct the pattern of the chimes out of pure tones at various places on the frequency spectrum.

Method

—Subjects—Four paid volunteers, one male and three female students at the University of Oregon, were employed as subjects. All four were *nonmusical* in the sense defined earlier. They all stated that they had heard the NBC chimes many times, and they were able, either initially or after preliminary practice, to produce the pattern accurately at its normal frequency level.

—Apparatus—Both of the RCA audio generators used earlier were now controlled by the subject, for frequency and also for loudness. The Grayson Stadler Twin oscillator, still in the experimenter's room, provided on each trial some anchor tone at a 40-db sensation level. The subject could listen, through the Koss Pro-4 headphones, to any of these three oscillators by pressing one of three button-switches mounted on a panel in front of him, between his two oscillators. The vernier stroboscope was used for all frequency readings, as before.

—Procedure—The general procedure, employed on every trial, was to give the subject an anchor tone, which was designated as either the first, second, or third tone of the sequence, whereupon his task was to adjust the other two tones to make the pattern sound like the NBC chimes.

Initially the subject was interviewed about his musical background (or lack thereof) and acquainted with the apparatus, particularly the oscillator controls. He was then asked to construct the chimes pattern, given the first tone at its normal frequency. Errors of two or more chromatic steps were not uncommon on this first attempt. The pattern was then set correctly by the experimenter; the subject was allowed to listen to it and immediately thereafter attempted to reconstruct it given an anchor tone. This alternation of demonstration and production was continued for 10 practice trials, by the end of which three subjects were consistently accurate within a chromatic step. The fourth subject, whose errors were greater, was given 10 more practice trials, in the course of which he attained a similar accuracy level. (Two prospective subjects never achieved this level of accuracy during the practice session and were dropped.) It should be clearly understood that the practice session involved *only* the standard, or untransposed, pattern, at the frequencies G = 2-8 = 196 Hz, E = 3-5 = 329.6 Hz, and C = 3-1 = 261.6 Hz. At no time during the experiment was the subject given any transposed exemplar of the pattern. However, his memory was refreshed at the beginning of each regular session by a single presentation and reconstruction of the standard pattern, as in the practice session; occasionally he made an error greater than a chromatic step and received a second practice trial.

In the experiment proper, the subject was required to construct the pattern at 14 different frequency levels: the standard, 3 levels below it, and 10 above it. Adjacent levels were half an octave apart. Thus half the transpositions were within the key of C (that is, an integral number of octaves from the standard) and therefore involved the possibility of 'octave generalization,' whereas alternate transpositions were to the key of F-sharp/G-flat, which of all keys is the one most remote from C by musical standards. Except at the lowest frequency level, each of the three tones was used twice as the anchor, yielding a total of six trials at each level. Because of apparatus limitations, only the second and third tones could be used as anchors at the lowest frequency level. Each session lasted about an hour and consisted of one trial at each frequency level. Order of presentation was random. The anchor tone was randomly selected for each trial without replacement, so that after three sessions every anchor tone had been used at every frequency level, with the exception noted. This procedure was replicated in the second three sessions.

The subject was, of course, told at the beginning of each trial which of the three tones the anchor represented. He was instructed to sound only one tone at a time (more specifically, to use a single finger in pressing buttons), and to play the tones always in their proper sequence. He was further instructed to keep the variable tones equal in loudness to the anchor as he adjusted their frequencies.

Results

The NBC chimes pattern, as normally sounded, consists of an ascending interval of nine chromatic steps followed by a descending interval of four steps. We now consider the sizes of these intervals as reproduced by the subjects at various frequency levels; also, the variability of the reproduced intervals as a function of frequency. From the six trials of a subject at each frequency level, a median and an average deviation about the median were calculated for each of the two intervals. (Position of anchor tone made no consistent difference and was ignored. Medians were used to minimize the weight of occasional 'wild' responses, but averaging log frequency settings, as in Experiment I, would have yielded practically identical results.) The means of these values, over the four subjects, are graphed in Figure 4. This average function represents the performance of individuals quite well: one subject was considerably more variable than the other three, but there were no important qualitative differences of the sort encountered in Experiment I.

The functions shown in Figure 4 are, for all practical purposes, horizontal straight lines. Quite simply: the chimes pattern was transposed on a musical scale, or in a log frequency medium. Intervals were consistently set a little too large; this constant-error tendency was reliable enough, but trivial in magnitude. Nothing remotely suggestive of a mel scale appears in the data.

Again, the variability results are informative. Variable error (in log frequency units) is fairly uniform over the lower frequen-

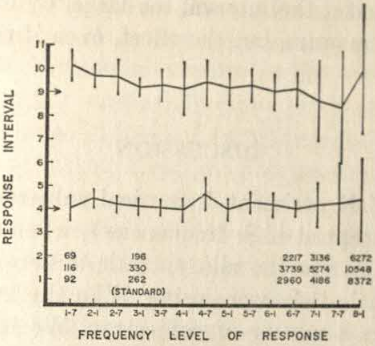


Fig. 4. The two intervals of the NBC chimes as reproduced at various frequency levels in Experiment II. Vertical lines extending ±1 mean deviation show variability. (Means were taken, over subjects, of within-subject medians and associated mean deviations.) Response intervals are measured in chromatic steps; the standard intervals were at 4 and 9 steps respectively. Adjacent frequency levels (abscissa) are separated by half an octave. Frequencies corresponding to perfect musical transposition of the pattern are specified at selected levels.

cies, up through level 6-7; see Figure 4. At level 7-1, where only the highest tone of the pattern exceeds 5,000 Hz, variability increases to more than double the average at lower frequencies. At the two highest levels—7-7 and 8-1, which lie entirely above 5,000 Hz but for one tone in the former—variability is almost four times the lower norm.

Thus, in terms of both scale and variability, we find the subjects of Experiment II—all of whom were 'nonmusical'—behaving essentially as did the 'musical' subjects in Experiment I.[3] A possible exception to this generalization is that variability in Experiment II was about 50% greater on the large interval than on the small one, whereas in Experiment I variability was dependent on interval size for nonmusical subjects but not for musical subjects. However, there is no certainty that this variability difference in Experiment II is attributable to interval size per se rather than to some other feature of the pattern—than to the fact that the second and smaller interval ends on the keynote or tonic, for example.

It is quite clear that 'octave generalization' was not an important factor in Experiment II. If it had been, we would presumably find an alternating pattern of variable error, low at the n-1 frequency levels and high at the n-7 levels. There is, in the function for the larger interval in Figure 4, some suggestion that *constant* *error* (tendency to set the interval too large, by musical standards) alternates in this manner, but the effect, even if real, is exceedingly small.

DISCUSSION

On the task of Experiment I, musical subjects transposed on a musical scale (except at high frequencies), whereas the behavior of nonmusical subjects was too idiosyncratic to permit any generalization about the scale they were using. *Why* the latter were so idiosyncratic remains a matter of conjecture. We have already questioned whether they were actually transposing patterns in the sense that our method presupposes. A further possibility is that their short-term memory traces (whatever the psychophysical process on which these traces were based) were highly unstable and hence vulnerable to a variety of disturbing influences. In any case, the results of Experiment II show that musically untutored persons cannot be supposed simply to lack the musical scale. Quite the contrary: they clearly transpose in a log frequency medium when given a well-defined, highly overlearned pattern as standard.

Whatever the mel may be, it is not a unit of melody; pitch as a medium has been much better described by the musician than by the psychophysicist. This is not to say that the mel scale has no psychological reality; musically competent observers, like Stumpf, have long recognized that musical intervals are larger (on some necessarily nonmusical scale) at high frequencies than at low frequencies. There have been numerous attempts to reconcile musical pitch (or 'tonality') with pitch in the mel-scale sense, often employing the spatial model of a tonal spiral. Boring's review of these efforts ends with this remarkable conclusion: "Nowadays, with diminution of the phenomenological interest in tone, with the failure to get differential limens of tonality, and with the knowledge that the aural harmonics insistently apply the musical relations to tones, the need for this attribute has largely evaporated. The octave is, so it is thought, the tonal distance for which the harmonics—and there are always aural harmonics—interpenetrate maximally" (1942, p. 380).

Our basic disagreement with Boring is perhaps clear without further belaboring, but the matter of harmonics does deserve discussion at this point. If one wishes to maintain that the musical scale is initially established in either the individual or the culture by relationships between harmonics, his position, though debatable, is a tenable one. However, if the argument is that some characteristic interpenetration of harmonics constitutes the necessary basis for identifying a particular musical interval, or is somehow essential to the encoding of tones in musical terms, the following considerations strongly disconfirm it. First, we find musical transposition of patterns that are entirely sequential, with no temporal overlap of tones. Second, tones were kept at a 40-db sensation level in both of the experiments reported—with close precision in Experiment I, approximately in Experiment II. At this level aural harmonics are not phenomenally evident, although they may be detected at the lower frequencies (where higher amplitudes are required for equal loudness) by the presence of beats with a probe tone. According to the curves that Fletcher (1929) obtained by the probe method, tones at the 40-db sensation level with frequencies greater than 500 Hz are entirely free of aural harmonics. And in our results the musical scale obtained for more than three octaves above 500 Hz. Third, the two musical subjects of Experiment I still made consistently musical transpositions when spot-checked at the uncomfortably low sensation level of 20 db. According to Fletcher's data, tones at this level are free of aural harmonics down to somewhat less than 120 Hz.

Perhaps the most provocative finding of these studies is the abrupt deterioration of musical transposition that occurred at about 5,000 Hz, both in the musical subjects of Experiment I and in the nonmusical subjects of Experiment II. Something changes rather dramatically at this level; phenomenally it is identifiable as a loss of musical quality, whatever that may be. The two musical subjects in Experiment I showed a break at the very same chromatic step. This is between three and four half tones above the high end of the piano; in other words, it would appear that just three keys could be added to the upper end of the piano keyboard without exceeding the musical range. The piccolo's highest note is about the same as the piano's, and no conventional instruments go higher; pipe organs have been constructed with notes as high as 12,000 Hz, but it is not clear that this extension of range is more than a curiosity, or an organ builder's tour de force. The empiricist may argue that tones above 5,000 Hz are unmusical because they do not occur as fundamentals in music, and the nativist may counter that they do not occur in music because they are unmusical; nothing in our data disproves either position.

A very interesting possibility, however, is that the mel scale is dependent on a place mechanism (as indeed a mapping of mel intervals on the basilar membrane strongly suggests) but that the musical scale is based on a frequency-volley mechanism with associated neural filters. It seems evident already that the place principle cannot account for very low frequencies, or the volley principle for very high ones, but the region and extent of overlap between the two mechanisms is ambiguous. The following of sound frequencies by squads of auditory fibers has been observed up to about 4,000 Hz (Wever, 1949), and it is not implausible to suppose a real upper limit near 5,000 Hz where we find so abrupt a break in transposition performance. The argument for this hypothesis would perhaps be more convincing if subjects definitely went over from a musical scale to a mel scale at 5,000 Hz (though it is not entirely clear what this would entail); what we actually find, particularly in Experiment II, is that the musical scale still obtains at higher frequencies on the average, albeit with much greater variability.

The failure of 'octave generalization' to manifest itself in Experiment II—that is, the fact that subjects transposed about as well from C to F-sharp as from C to C in different octaves—tells us something about the terms in which the pattern was stored. Bear in mind that in this experiment the transposition was from a standard held in long-term memory: if the normal chimes pattern had been presented before each trial, octave relationships might well have assumed greater importance. (Experiment I did employ short-term memory but provided no direct opportunity for octave generalization.) The circulating short-term memory of a tonal sequence just heard, which is experienced as an auditory image extended in real time (as a melody that 'runs through one's head'), typically preserves the key or specific pitch values of the original. On the other hand, a subjectively similar image generated from long-term memory is no less determinate in key,[4] but its key may or may not be that of the original. (If the reader can form a clear image of the NBC chimes, he can then go to the piano and find the three particular notes that match his image, in pitch as well as in pattern, to the nearest half tone. It is quite unlikely, however, that the key of his image—identified by the final note—will be that of C, in which he has always heard the chimes.) The long-term memory trace is encoded in terms based upon relations, or intervals, or distances on a log frequency scale. There is nothing about such a coding system that would preclude the additional or auxiliary storage of information about specific pitches, or key,[5] but it is evident that normal individuals do not, in fact, preserve this kind of information with any high degree of precision. If they did, 'absolute pitch' would be commonplace.

As suggested at the beginning, pitch is only one of a number of psychophysical continua that may legitimately be studied as media. In other, fairly obvious ways, pitch is unique, and the present results, though enlightening in the case of pitch, do not tell us what to expect more generally. It will be interesting to try to quantify the morphophoric properties of other continua such as loudness, but avoiding pitfalls of the kind that we encountered with nonmusical subjects in Experiment I may require considerable care and ingenuity.

Notes

The research was supported in part by the Air Force Office of Scientific Research (Grant 973-66) and in part by the Advanced Research Projects Agency of the Department of Defense (Contract F44620-67C-0099). Dr. Olson is now at the University of Colorado. Received for publication August 6, 1970.

1. Strictly speaking, the midpoints of odd-numbered intervals were half a chromatic step (one twenty-fourth of an octave) higher than those of even-numbered intervals.

2. Note, however, that that variability is measured in chromatic steps, or log frequency units. In terms of frequency, what is averaged over the lower octaves is, in effect, *proportional* variability, with interval size held constant.

3. One of the more bizarre patterns obtained in Experiment I was that of the nonmusical subject LC (Figures 1, 2, 3). After the apparatus was rearranged for Experiment II, LC was brought back to the laboratory, and made a number of transpositions of the NBC chimes. Although the conditions under which she was run were different in detail from those later settled upon for Experiment II, her behavior on the new task was, in all essential respects, identical with that of the regular subjects in Experiment II.

4. Why the image has to be in any key, when the memory trace presumably is not, is an interesting question (see Attneave, 1967).

5. Whether persons with 'absolute pitch' actually have stable, long-term representations for specific pitches or whether they represent key (tonality) in some different manner is quite obscure. Apparently there are individuals who can correctly identify a tone as a C, an F, or whatever, but nevertheless make octave errors (see Bachem, 1937).

References

Attneave, F. 1967. Criteria for a tenable theory of form perception. In *Models for the perception of speech and visual form,* ed. W. Wathen-Dunn, pp. 56–67. Cambridge, Mass.: The M.I.T. Press.

Bachem, A. 1937. Various types of absolute pitch. *Journal of the Acoustical Society of America* 9:146–151.

Boring, E. G. 1942. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century.

Fletcher, H. 1929. *Speech and hearing.* New York: Van Nostrand.

Stevens, S. S., and Volkmann, J. 1940. The relation of pitch to frequency. *American Journal of Psychology* 53:329–353.

Stevens, S. S., Volkmann, J., and Newman, E. G. 1937. A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America* 8:185–190.

Wever, E. G. 1949. *Theory of hearing.* New York: Wiley.

DISTRIBUTED PRACTICE IN PAIRED-ASSOCIATE LEARNING: INTERSTIMULUS SIMILARITY

Mukul K. Dey
McMaster University, Canada

The subjects learned CCC trigrams at five levels of intralist stimulus similarity and at two intertrial intervals (2 and 60 sec.). There was a significant interaction of similarity and interval: the greater the similarity, the greater the facilitative effect of the 60-sec interval. This is interpreted as support for the theory that interference in paired-associate learning is due to the conflict of simultaneously evoked response tendencies.

A reduction in the level of interference through the rest intervals is often considered to be the basis of facilitation of paired-associate learning by distribution of practice. Two interpretations of the nature of this interference have been offered. One of these interpretations (Underwood, 1961) attributes interference to the response-learning phase. After a careful review of pertinent findings, Underwood and Ekstrand (1967) have recently rejected this proposition. The other interpretation, suggested by Underwood and Schulz (1959, 1961), links interference to the associative processes. The writer has developed this last-mentioned idea into a theory that conceives interference in paired-associate learning as a temporally labile inhibitory potential generated by the conflict of simultaneously evoked response tendencies (Dey, 1966).

A principal prediction from the theory thus developed is that distribution of practice—as either intertrial or interpair interval—and interstimulus similarity should influence learning in an interactive pattern. This prediction has been partly substantiated by the demonstration of a joint influence of interstimulus similarity and interpair interval: with higher values of the former, an increase in the latter leads to greater facilitation (Dey, 1966). However, for a complete test of the prediction, it is necessary to determine if a similar relationship prevails when intertrial interval is used as the medium of variation. An experiment by Marshall and Runquist (1962) failed to show such a relationship, though the main effects of both intertrial interval and interstimulus similarity occurred in significant proportions. These authors used CVC trigrams as stimuli and arranged two levels of interstimulus similarity. The experiment reported here sought further evidence, using CCC trigrams as stimuli and varying interstimulus similarity through a much wider range.

METHOD

—Subjects—The subjects were 180 sophomores and juniors from the writer's classes. They were randomly assigned to ten groups of equal size. Although none was lost for failure to learn, 7 subjects were replaced because they either gave evidence of having misunderstood the instructions or reported inadequate interest or motivation. The subjects' mean age was 23.4 years, in the range from 18 to 31 years. The male-female ratio was 3:2 in every group.

—Design and material—The experiment was given a 2 × 5 factorial design, with intertrial interval at 2 or 60 sec and interstimulus similarity varied in five ways by use of five series of six trigrams as stimuli (Table 1). The trigrams in series A through E, representing a progressive increase of interstimulus similarity, were constructed with 18, 13, 9, 6, and 3 consonants respectively. The three-letter words adopted as response terms in all lists are *ham, fig, ray, box, jet,* and *cup*. Rotating the factor of pairing so that each stimulus was paired once with every response term, six lists of paired associates were prepared for each degree of interstimulus similarity.

—Apparatus and procedure—The auditory stimuli were presented by means of a Tandberg tape recorder with solenoid start-stop mechanism. The response terms were visually presented through projection on a table-view screen by a Leitz Pradovit-Color slide projector with a remote-control and automatic-return mechanism. A Scientific Prototype sound-operated relay, three Hunter interval timers and two Lehigh Valley Electronics pulseformers assisted in

Table 1. The five series of trigrams representing five degrees of interstimulus similarity in increasing similarity

A	B	C	D	E
HLF	HFB	HFB	HFB	GKT
BJX	BXJ	FXJ	BXJ	GTK
SQZ	XQZ	BQZ	JQH	TKG
RMY	RMQ	RHM	XBF	TGK
NDC	NDR	JZR	QJX	KTG
VWP	VNP	QMX	FHQ	KGT

controlling the temporal variables. The stimuli appeared at intervals of 9 sec: 4 sec after the start of every stimulus presentation, the response term was exposed for 2.5 sec. Assuming that the time consumed in the presentation of an entire stimulus trigram was about 1.5 sec, the anticipation interval was roughly 2.5 sec, and making some allowance for the relay times, the interpair interval may be estimated as 2 sec. Three 60-min audio tapes were prepared for every stimulus series, repeatedly recording three 24-item sequences, one on each tape. The three stimulus sequences were arranged in such a manner that the use of each in conjunction with the same six sequences of 24 slides—4 for each word—would furnish the desired rotation of the factor of pairing.

The instructions were those normally used in paired-associate learning by the anticipation method, plus some explanatory remarks on the rest-interval activity. For all subjects, practice was carried on to the criterion of first correct elicitation of all responses on a trial. The groups receiving 2-sec intertrial intervals practiced continuously, as in their case intertrial and interpair intervals were identical. The groups receiving 60-sec intertrial intervals discriminated between pure tones spaced an octave apart as the rest-interval activity. A second Tandberg tape recorder with solenoid switch, which presented the pairs of tones at intervals of 9 sec, was started at the moment the first tape recorder (delivering stimuli) was stopped at the end of every trial. The two tones in each pair, which were of 1-sec duration each, were separated by a .5-sec interval. The subjects reported their judgments with reference to the second tone. Six judgments were called for during each 60-sec intertrial interval.

RESULTS

An analysis of variance of the group means for the number of trials to criterion (Table 2) revealed significance both in progressive slowing-down of learning with increased interstimulus similarity [$F(4, 170) = 18.2, p < .01$] and in facilitative influence of distribution of practice as evidenced in quicker learning with 60-sec than with 2-sec intertrial intervals [$F(1, 170) = 6.82, p < .01$]. Further,

Table 2. Group means of number of trials to criterion and of number of overt errors per trial

Stimulus series	Trials to criterion				Overt errors per trial			
	2 sec		60 sec		2 sec		60 sec	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
A	15.50	5.53	15.72	4.55	1.57	.87	1.45	.75
B	16.28	5.12	18.06	4.58	1.66	.76	1.75	.71
C	20.94	5.53	18.78	6.14	1.89	.67	1.64	.66
D	24.50	6.29	17.83	4.18	2.09	.98	1.88	.69
E	30.89	10.48	24.78	9.60	2.27	.94	2.12	.83

the magnitude of the intertrial-interval effect clearly seems to have increased with increased interstimulus similarity. This finding was proven reliable by a significant interaction variance [$F(4, 170) = 2.86, p < .05$].

For each subject, the total number of overt errors, most of which consisted of intralist intrusions, was divided by the number of trials to learn. An analysis of variance of the group means for these data (Table 2) showed a positive relation of error frequency to interstimulus similarity [$F(4, 170) = 4.23, p < .01$] but an indeterminate effect of intertrial interval [$F(1, 170) = 1.24$, $\mathrm{p} > .25$], as well as lack of interaction between the two variables [$F(4, 170) < 1$].

DISCUSSION

A theory previously advanced by the writer (Dey, 1966) assumes the development of an inhibitory state as an inevitable product of associative growth and accounts for distribution-of-practice facilitation by postulating a dissipating property in this inhibitory state. A brief account of this theory follows. During the performance of any list component, (*a*) the excitatory potential pertinent to the correct response is released and (*b*) the responses belonging to the other list components are evoked by stimulus generalization. The conflict from these simultaneously incompatible response tendencies results in the development of an inhibitory potential. During the interpolated rest intervals, both excitatory and inhibitory potentials are dissipated, but since the dissipation rate is faster for the latter than for the former, the net effect of the rest intervals is some degree of facilitation. By assuming a positive relationship between the conflict-generated inhibition and the strength of conflicting responses, which in turn is a positive function of stimulus similarity, this theory credits a critical role to interstimulus similarity in the distribution-of-practice facilitation of paired-associate learning. The evidence of the interaction of interstimulus similarity with interpair interval in a previous study (Dey, 1966) and with intertrial interval in the present study upholds this point.

The positive relationship between error frequency and interstimulus similarity, as shown in the present experiment and in certain previous studies (Feldman and Underwood, 1957; Underwood, 1951, 1953a), conforms to the theory here proposed. It may be argued that with decreased similarity between stimuli, the strength of the generalized responses is increasingly closer to the reaction threshold and hence the probability of their overt elicitation is progressively less. On the other hand, the present experiment has shown a neutral influence of the distribution-of-practice variable upon error frequency. This is as would be expected, since the theory explains reduction in the level of interference through rest intervals by postulating differential dissipation of excitatory and inhibitory potentials rather than differential forgetting of correct and incorrect responses and, therefore, does not predict any difference in the strength of the generalized responses when either interpair or intertrial interval is manipulated.

Although the theory that attributes distribution-of-practice facilitation to interference in response learning does not stand on a firm ground, the importance of response learning in the total learning process cannot be overlooked. Large positive transfer to paired-associate learning from the prior free learning of response terms (Underwood, Ekstrand, and Keppel, 1964; Underwood, Runquist, and Schulz, 1959) and the positive relation of such transfer to the degree of prior free learning (Underwood and Ekstrand, 1967) prove the importance of response learning. It is also to be acknowledged that in certain circumstances some kind of interference may indeed be generated in the course of response learning. Using such response terms as are composed of discrete elements, several studies have shown considerable list difficulty as a result of increase in the proportion of elements that are common between the response terms (Levitt and Goss, 1961; Newman and Buckout, 1962; Rubin and Brown, 1967; Underwood, Ekstrand, and Keppel, 1964). Such findings would not be obtained unless some kind of interference was produced while the elements in the response terms were being integrated.

Notes

Reprints are available from the author, whose mailing address is P.O. Box 36, McMaster University, Hamilton, Ontario, Canada. Received for publication July, 1970.

References

Dey, M. K. 1966. An explanation of distributed practice efficacy in paired-associate learning. *Journal of General Psychology* 74:61–75.

Feldman, S. M., and Underwood, B. J. 1957. Stimulus recall following paired-associate learning. *Journal of Experimental Psychology* 53:11-15.

Levitt, H., and Goss, A. E. 1961. Stimulus attributes and drive in paired-associate learning. *Journal of Experimental Psychology* 62:243-252.

Marshall, M. A., and Runquist, W. N. 1962. Facilitation in paired-associate learning by distributed practice. *Journal of Verbal Learning and Verbal Behavior* 1:258-263.

Newman, S. E., and Buckout, R. 1962. S-R and R-S learning as functions of intralist similarity. *American Journal of Psychology* 75:429-436.

Rubin, E. D., and Brown, S. C. 1967. Constant versus varied serial order in paired-associate learning: The effect of formal intralist similarity. *Journal of Experimental Psychology* 73:257-267.

Underwood, B. J. 1951. Studies of distributed practice: II. Learning and retention of paired-adjective lists with two levels of intralist similarity. *Journal of Experimental Psychology* 42:153-161.

Underwood, B. J. 1953a. Studies of distributed practice: VIII. Learning and retention of paired nonsense syllables as a function of intralist similarity. *Journal of Experimental Psychology* 45:133-142.

Underwood, B. J. 1953b. Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. *Journal of Experimental Psychology* 45:143-149.

Underwood, B. J. 1961. Ten years of massed practice on distributed practice. *Psychological Review* 68:229-247.

Underwood, B. J., and Ekstrand, B. R. 1967. Effects of distributed practice on paired-associate learning. *Journal of Experimental Psychology Monograph Supplement* 73(4, pt. 2, whole no. 634).

Underwood, B. J., Ekstrand, B. R., and Keppel, G. 1964. Studies of distributed practice: XXIII. Variations in response-term interference. *Journal of Experimental Psychology* 68:201-212.

Underwood, B. J., Runquist, W. N., and Schulz, R. W. 1959. Response learning in paired-associate lists as a function of intralist similarity. *Journal of Experimental Psychology* 58:70-78.

Underwood, B. J., and Schulz, R. W. 1959. Studies of distributed practice: XIX. The influence of intralist similarity with lists of low meaningfulness. *Journal of Experimental Psychology* 58:106-110.

Underwood, B. J., and Schulz, R. W. 1961. Studies of distributed practice: XX. Sources of interference associated with differences in learning and retention. *Journal of Experimental Psychology* 61:228-235.

SELECTIVE PERCEPTION AND THE NUMBER OF ALTERNATIVES

Kent Gummerman
University of Texas at Austin

The subjects saw one of 4 or one of 16 letterlike stimuli, with the choices reduced to 2 before and after (groups BA4 and BA16) or after (groups A4 and A16) the stimulus, which was followed by a masking figure. Identification accuracy was the same for groups BA4, A4, and BA16; it was considerably poorer for group A16. The outcome suggests that selective perception occurs with prior knowledge of alternatives—but only if that knowledge simplifies the task.

Lawrence and Coles (1954) were among the first to use a 'before-after' method to study perceptual tuning (selective perception). In their paradigm the subjects learn the stimulus alternatives either before or immediately after the presentation of a tachistoscopically presented stimulus. If the stimulus is identified more accurately with foreknowledge of the alternatives than without, the explanation may lie in the selective tuning or preparation of perceptual functioning that foreknowledge can allow.

Lawrence and Coles found that prior knowledge of alternatives did not enhance performance, but they presented the alternatives as written words rather than pictorially: their stimuli were pictures of objects, and it is unlikely that verbal alternatives could have provided enough information about the stimulus pictures to allow the subjects to prepare for selective processing of critical stimulus features. Egeth and Smith (1967) noted this potential problem and used pictorial alternatives in a similar before-after experiment. They also followed the suggestion of Long, Reid, and Henneman (1960) and presented the alternatives either (*a*) after the stimulus or (*b*) both before *and after* the stimulus; the latter condition relieves the memory load that exists when alternatives are shown only before the stimulus. Egeth and Smith found that foreknowledge facilitated identification, though this facilitation might have been caused by selective ocular orientation prior to the stimulus presentation rather than by selective perceptual functioning (Gummerman and Munsinger, 1970). With alternatives known before the stimulus, the subjects could have fixated directly on critical areas of the stimulus field; without prior knowledge of the alternatives, fixation could only have been at a neutral point, leaving critical features falling on less sensitive retinal locations.

Gummerman (1970) used a before-after design with pictorial alternatives that made selective fixation impossible, and still he found large facilitation with foreknowledge; however, when a patterned masking figure was introduced immediately after the stimulus, the effect vanished. When no mask was used, the subjects with prior knowledge of alternatives apparently were able to make full use of the fading stimulus icon (Averbach and Coriell, 1961; Sperling, 1960; Neisser, 1967). When the alternatives were withheld until immediately after the stimulus, the icon of the stimulus was fading rapidly while the alternatives were being processed (see Eriksen and Collins, 1969) and was of little use after the alternatives were known. When an effective mask followed the stimulus, processing of the information in the icon was forced to occur before the physical stimulus terminated (i.e., before the mask appeared) in all conditions. The mask ensured that performance scores were unconfounded measures of perceptual processing, so the equality of performance with and without foreknowledge was strong evidence against the selective-perception notion.

One aspect of the Gummerman (1970) study limits the generality of the conclusion just stated: only four stimuli were used, and knowledge of alternatives consisted of reducing the choices from four to two. In essence, then, the study compared processing when the stimulus was either one of two possibilities (the foreknowledge condition) or one of four possibilities (no foreknowledge). It is conceivable that optimally efficient processing strategies were very similar in these two cases, so that foreknowledge did not significantly simplify the function required for stimulus identification. If a considerably larger number of stimuli had been used, foreknowledge might have facilitated identification.

To explore this possibility, the present study used two groups of subjects in a before-after design, with a mask following the stimulus in all conditions. One group saw only 4 stimuli (as in the Gummerman, 1970, study), with the choices reduced to 2 either before and after or after the stimulus; but the other group saw 16 stimuli in all, with choices reduced to 2 either before and after or after the stimulus. Foreknowledge was not expected to facilitate performance for the first group; but, assuming that the concept of selective perception is valid, foreknowledge should have a large positive effect for the subjects dealing with 16 stimuli. On the other hand, if perceptual functioning cannot be modified in any way by more precise specification of the stimulus to be presented, then the performance of all subjects should be identical.

METHOD

—Subjects—The subjects were 32 undergraduate females from introductory psychology courses at the University of Texas at Austin. All had normal vision, either with or without correction, and had had no previous experience in perceptual experiments. Each subject was assigned randomly to one of four groups: group BA4 (4 stimuli, 2 alternatives shown before and after the stimulus), group A4 (4 stimuli, 2 alternatives shown only after the stimulus), group BA16 (16 stimuli, 2 alternatives shown before and after the stimulus), or group A16 (16 stimuli, 2 alternatives shown only after the stimulus).

—Apparatus and stimulus materials—A three-field tachistoscope (Scientific Prototype, GB), with a 104-cm viewing distance, was used to present all displays. The luminances of all displays were about 20 ftL for white portions and about 1 ftL for black portions.

All figures were black on white backgrounds. The stimuli were 4 × 4 square matrices, with 12 cells blackened to form likenesses of the capital letters A, C, F, G, H, I, K, L, N, O, P, R, T, U, Y, and Z. (Letterlike forms were used to provide convenient labels for the subjects.) The matrices, drawn with black ink on white index cards, subtended 40 min visual angle on each side. A black fixation point, 5 min in diameter, appeared at the center of the display. To prevent selective fixation on a particular area of the stimulus field, the stimulus could appear with equal likelihood in one of three positions: 8 o'clock, 12 o'clock, and 4 o'clock on an imaginary clock face centered on the fixation point. The distance from the fixation point to the nearest part of each stimulus location subtended 20 min.

The masking pattern—which always immediately followed the termination of the stimulus exposure—was a 60 × 60 min black-and-white 'checkerboard,' six small squares on a side (each small square subtending 10 min visual angle, the same as the cells of the stimulus matrices). Twenty-two of the cells were selected randomly and blackened. (When superimposed over the stimulus letters, the mask shared 11 identical cells with letters C, H, and I; with all other letters, 9 cells were identical.) The complete mask display consisted of three identical checkerboards: one of the checkerboards was precisely positioned over the center of each of the three stimulus locations.

The alternatives were the same size as the stimuli and appeared side-by-side (separated by 40 min) 80 min above the top of the 12 o'clock position.

The subjects in groups BA16 and A16 saw each stimulus three times (once in each stimulus position). On each of these 48 trials, a different pair of alternatives was displayed; each of the 16 stimulus letters appeared in a pair of alternatives six times (three times on the left and three times on the right).

The subjects in groups BA4 and A4 were divided into four subgroups according to which 4 of the 16 stimuli were presented to them: ALNT, CFKR, GHIP, or OUYZ. (These groupings were made randomly.) In each subgroup the subjects saw each stimulus a dozen times (four times in each stimulus position) and each of the six possible pairs of alternatives eight times.

In A4 and A16, two 'dummy' alternatives were shown before the stimulus to make the visual displays similar to those of BA4 and BA16. The dummy alternatives were the same size as the real ones and appeared in the same location. Each dummy had 12 cells blackened, forming a 'plus.'

—Procedure—At the beginning of each session, the experimenter explained the task and allowed the subject to examine each stimulus for several seconds. Thus each subject was aware of the size and composition of the set of stimuli to be used. Sixteen practice trials were then taken in the tachistoscope, after which the 48 data trials were begun. A 5-min rest period was taken after the first 24 trials.

For each subject the following procedure was used. A trial began with an 8-sec presentation of two alternatives and the fixation point. (The alternatives were 'real' in BA4 and BA16. The 'dummy' alternatives were shown in A4 and A16; the subjects were told to look at them as long as they would have if they had been real.) The subject examined the alternatives and then fixated the central point. At the offset of the alternatives, the stimulus automatically appeared (in one of the three possible positions) for 120 msec. At the offset of the stimulus, the alternatives appeared (groups A4 and A16) or reappeared (groups BA4 and BA16) for 8 sec, accompanied by the mask patterns. The subject responded by stating the name of the letter she believed was presented; she responded whenever ready, which was usually about 3–5 sec after the stimulus offset. A response was required on every trial; no trial-by-trial feedback was given. The intertrial interval (during which the display was black) was about 2 sec for groups BA16, A4, and A16; the interval was about 4 sec for group BA4. (The BA4 intertrial interval was longer than the others because the displays required more time to change.) A full session required about three-fourths of an hour.

RESULTS AND DISCUSSION

Table 1 provides all data, averaged over subjects.[1] While performance in the groups that had foreknowledge of the alternatives (the BA groups) was significantly better than in the groups that did not have such foreknowledge (the A groups) [$F(1, 28) = 5.98$, $p < .02$], the interaction between the number of stimuli and foreknowledge, shown in the last line of the table, is the critical effect for this study. Al-

Table 1. Percent correct identifications

Stimulus location	Group BA4	Group A4	Group BA16	Group A16
8 o'clock	83.6	89.1	81.3	73.4
12 o'clock	90.6	78.1	91.4	71.9
4 o'clock	79.7	86.7	92.2	73.4
Average over stimulus location	84.6	84.6	88.3	72.9

Note: Averages for 8 o'clock, 12 o'clock, and 4 o'clock are 81.9, 83.0, and 83.0 respectively.

though performance with or without foreknowledge was identical for the subjects presented only 4 stimuli (confirming the Gummerman, 1970, finding), performance with foreknowledge was considerably better than performance without for the subjects seeing 16 stimuli. This interaction effect was reliable $[F(1, 28) = 5.98, p < .02]$.[2] This outcome is clear evidence that foreknowledge of alternatives can facilitate identification of tachistoscopic stimuli—but only when foreknowledge significantly reduces the size of the effective stimulus set (the number of stimuli that the subject must consider at the time of stimulus exposure).

When the total stimulus set contained only 4 stimuli, reducing the effective set to 2 stimuli apparently allowed no noticeable change in perceptual processing strategies. In other words, perceiving one of 4 stimuli was no more difficult than perceiving one of 2. However, when the effective set contained 16 stimuli, the perceptual task was more demanding. And while group A16 may have extracted no less information from the stimulus than any other group, the probability was lower that the information the subjects obtained was precisely the information they needed to discriminate between the members of all the pairs that could have appeared after the stimulus. Foreknowledge made it more likely, for group BA16, that the information obtained was relevant to the one pair in question. Selective perception, then, does exist; it appears to consist of 'instructions' as to which of the stimulus attributes are task-relevant and should be processed.

Two further outcomes should be mentioned. First, the size of the difference between the groups that had foreknowledge and those that did not was not the same in all three stimulus locations $[F(2, 56) = 4.45, p < .02]$. The difference was greatest at the 12 o'clock and smallest at the 8 o'clock position. The cause of this effect is not clear; but since it did not interact further with the number of stimuli, the principal outcome of the study is not affected.

Second, the performance of subjects who saw only four stimuli differed according to the particular subsets of stimuli presented them. The average scores corresponding to the subsets were: ALNT, 95.8%; CFKR, 82.3%; GHIP, 87.5%; OUYZ, 72.9%. In an analysis of variance done on the scores of the subjects who saw four stimuli, this subset effect was significant [$F(3, 8) = 4.50$, $p < .05$] but did not interact with foreknowledge of the number of alternatives or with stimulus location. A subset difference had been foreseen, but it was expected to be related to the number of identical cells in the members of the six alternative pairs in each subset. No such relationship resulted, however; and no other reason for these differences is obvious.

Notes

This research was supported by the National Science Foundation's University Science Development Program, Grant GU-1598. Thanks are extended to Richard C. Kevin for preparation of the apparatus and stimulus materials; to George A. Hill for data collection and analyses; and to Clarke A. Burnham, Dennis McFadden, and Cynthia R. Roberts for their comments on an early version of the manuscript. The author's address is at the Department of Psychology, Mezes Hall 211, University of Texas at Austin, Austin, Texas 78712. Received for publication August 12, 1970.

1. Four subjects were discarded and replaced. Three were in group A16: one had previous tachistoscopic experience in a speed-reading course, one had had no sleep the night before, and one was extremely sensitive to bright lights. One subject in BA4 was receiving medication that dilated the pupils. If these subjects had not been replaced, the average A16 score would have been 70.1% correct and the average BA4 score 84.4% correct. The results of the statistical analyses would not have been different.

2. Coincidentally, the mean squares for this interaction (foreknowledge by number of stimuli) and for the main effect of foreknowledge (BA versus A) were identical.

References

Averbach, E., and Coriell, A. S. 1961. Short-term memory in vision. *Bell System Technical Journal* 40:309–328.

Egeth, H., and Smith, E. E. 1967. Perceptual selectivity in a visual recognition task. *Journal of Experimental Psychology* 74:543–549.

Eriksen, C. W., and Collins, J. F. 1969. Temporal course of selective attention. *Journal of Experimental Psychology* 80:254–261.

Gummerman, K. 1970. Knowledge of alternatives and perception of tachistoscopic stimuli. *Journal of Experimental Psychology* 83:385–390.

Gummerman, K., and Munsinger, H. 1970. Selective mechanisms in visual perception. *Perception and Psychophysics* 7:51–56.

Lawrence, D. H., and Coles, G. R. 1954. Accuracy of recognition with alternatives before and after the stimulus. *Journal of Experimental Psychology* 47:208–214.

Long, E. R., Reid, L. S., and Henneman, R. H. 1960. An experimental analysis of set: Variables influencing the identification of ambiguous, visual stimulus-objects. *American Journal of Psychology* 73:553–562.

Neisser, U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.

Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs* 74(11, whole no. 498).

SUM OF EVENT PROBABILITIES AND SUBJECTIVE ESTIMATES OF PI IN PROBABILITY LEARNING

Gloria J. Fischer
Washington State University

Students had 300 prediction trials on two-choice contingent-probability event ratios with sums less than, equal to, or greater than one; then they estimated the probability of the most frequent event. Their asymptotic prediction varied with their subjective estimates of the event's probability, rather than with its actual probability of occurrence. Specifically, they underestimated when the ratio summed to less than one and overestimated when it summed to more than one, apparently because they assumed that the events were mutually exclusive.

Human subjects usually match event frequencies in their environment; that is, they predict events in accordance with their frequency of occurrence. Only rarely is the most frequently occurring event, E_1, predicted more often than its probability of occurrence, π. This *overshooting* is produced by an increase in the utility of predicting correctly or incorrectly (Siegel, 1959). It also occurs with the use of a procedure designed to convince the subjects that the sequence of events is random, as in a gambling task (Goodnow, 1955; Peterson and Uhlela, 1965). Studies varying π differences and π ratios suggest that overshooting also occurs when the ratio sums to less than one ($\Sigma\ \pi_K < 1$, where π_K is the probability of event $_K$).

Such π ratios with sums of less than one have been used in several studies of contingent-probability learning (e.g., Brand, Sakoda, and Woods, 1957; Brand, Woods, and Sakoda, 1956). In 28 of some 30 such instances, the proportion of E_1 predictions, $p(A_1)$, was greater than the actual probability of that most frequently occurring event; in brief, overshooting occurred. The majority of cases occurred in the two studies cited, both of which varied the difference between two-event probabilities over a wide range of π ratios [π differences $(\pi_1 - \pi_2) = .2, .3, .4,$ or $.5$, and π ratios $= 2, 3, 6,$ or 11]. An increase in π difference (especially when the difference was greater than or equal to .4), but not in π ratio, was found to enhance $p(A_1)$.

Several studies have varied π ratios that did sum to one ($\Sigma\, \pi_\kappa = 1$); and matching, rather than overshooting, has been found. One of the most elaborate of these studies used a three-choice task and decreased π systematically (Little, Brackbill, Isaacs, and Smelkinson, 1963). The $p(A_1)$ was found to vary directly with π, with no evidence of overmatching at low values of π. On the contrary, there was a tendency to undermatch π values less than or equal to .7. This latter finding suggests that overshooting of π when the ratio sums to less than one ($\Sigma\, \pi_\kappa < 1$) is not due to low values of π. It might be due, rather, to some effect of blank (E_0) trials on the subjects' subjective estimates of event probabilities. That is, the subjects might tend merely to fill the void—to distribute the E_0 trials equally among the events and thus estimate a 65:15 ratio as 75:25. This is especially likely if the subjects are told that one of two possible predictions is correct on every trial, as was done in many of the contingent-probability learning studies cited (e.g., Brand, Sakoda, and Woods, 1957). If the subjects distributed E_0 trials proportionally rather than equally in a 65:15 ratio, their estimate would be closer to 80:20. Either way, overshooting would be substantial.

The present experiment tested further the hypothesis that ratio sums of less than one enhance the proportion of E_1 responses, and attempted to relate the effect to subjective estimates of event probabilities. The hypothesis that $\Sigma\, \pi_\kappa < 1$ enhances $p(A_1)$ was tested by varying π ratios so that $\Sigma\, \pi_\kappa > 1, = 1,$ or < 1, while holding the π difference a constant .5. Groups were given 300 trials on choice ratios of 95:45, 85:35, 75:25, 65:15, 55:5, or 50:0. If $p(A_1)$ varies with π, the relation between $p(A_1)$ and π should be essentially linear, with slope π. If $\Sigma\, \pi_\kappa < 1$ enhances $p(A_1)$, however, the relation between $p(A_1)$ and π or $\Sigma\, \pi_\kappa$ should be curvilinear, with slope less than π. The possibility that overshooting of π with $\Sigma\, \pi_\kappa < 1$ results from effects of E_0 trials on subjective estimates of event probabilities was assessed by asking the subjects to estimate π at the end of the experiment.

METHOD

—Subjects—A total of 102 students from introductory psychology courses were assigned randomly to one of three experimenters and to one of six π ratios. Participation fulfilled part of a course requirement.

—Apparatus—A large white board facing the subject contained a white signal light centered above two red event lights. Immediately below the event lights were two predictor levers. Each lever was connected to a preprogrammed punched tape with the appropriate probability of triggering its respective event light. A timer was connected so that in a 15-sec interval after a lever press, the tape controlled by that lever advanced one step. Eleven tapes were punched so as to yield π values over 300 trials of .95, .85, .75, .65, .50, .45, .35, .25, .15, or .05. The sequence of events was randomized independently for each π value (for each tape) over successive sets of 100-trial blocks. It can be seen, therefore, that the six π ratios used (50:0, 55:5, 65:15, 75:25, 85:35, and 95:45) were derived from these 11 tapes.

—Procedure—The subjects were instructed to predict the light they thought would come on by pressing the lever immediately below it. They were told to avoid a trial-by-trial approach as much as possible and, instead, to use a strategy that would yield as many correct guesses as possible over the entire session of 300 trials. The subjects got feedback on correct prediction trials only, thus learning under a noncorrection or contingent method. At the end of 300 trials they were asked to estimate the probability of the left (π) event.

RESULTS AND DISCUSSION

Figure 1 shows group mean proportions of E_1 predictions, $p(A_1)$, over the last block of 100 trials, as a function of the actual E_1 event probabilities, π. Shown too are the mean subjective estimates of E_1 probabilities, $\hat{\pi}$. It is clear, first of all, that the subjects' predictions did not vary as a function of π, since the slope of $p(A_1)$ deviates substantially from π. Instead, as expected from results of prior studies, $p(A_1)$ undershot π when $\Sigma\ \pi_\kappa > 1$ and overshot π when $\Sigma\ \pi_\kappa < 1$. Alone, these results support the view that ratio sums of less than one produce overmatching. However, the similar shape and slope of the $p(A_1)$ and $\hat{\pi}$ curves shows that the subjects actually were overmatching their own subjective estimates of π. Since this overmatching occurred at all π ratios, it cannot be attributed merely to ratio sums of less than one ($\Sigma\ \pi_\kappa < 1$). It is likely to have resulted, instead, from the instructions used, since they were based on instructions previously found to produce overmatching (Braveman and Fischer, 1968).

If prediction, $p(A_1)$, varies directly with subjective estimate, $\hat{\pi}$, why doesn't the estimate vary directly with the actual probability, π? A comparison of $\hat{\pi}$ and π curves in Figure 1 shows, instead, that

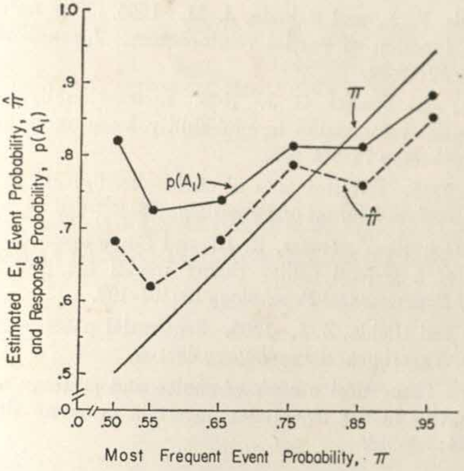


Fig. 1. The relation between the proportion of trials 201–300 on which subjects predicted the most frequent event, $p(A_1)$; the subjects' subjective estimate of the most frequent event probability, $\hat{\pi}$; and the actual probability of the most frequent event, π

$\hat{\pi}$ is above π when $\Sigma\ \pi_\kappa < 1$ and below π when $\Sigma\ \pi_\kappa > 1$. Clearly, the subjects must have assumed that the events were mutually exclusive—even though they were not told this was so, and in spite of some extremely deviant (from $\Sigma\ \pi_\kappa = 1$) π ratios, such as 50:0 and 95:45. Under- and overestimation of π in such circumstances resulted necessarily from the mutual exclusivity the subjects imposed on the sequence of choice events. One may conclude that $p(A_1)$ varies with $\hat{\pi}$ in human probability learning. And that $\hat{\pi}$ in turn varies with π, *except* when events are not mutually exclusive (as in contingent-probability learning) and deviate substantially from the ratio sum ($\Sigma\ \pi_\kappa = 1$) imposed by the subjects' assumptions that the events are mutually exclusive.

Notes

The author thanks several undergraduate students for their assistance: John Chambers and Steve Start helped with construction of the apparatus, and they and Sally Pederson collected pilot data; Joe Casey, Dave Tice, and Diane Willoughby collected the data reported. Received for publication August 3, 1970.

References

Brand, H., Sakoda, J. M., and Woods, P. J. 1957. Contingent partial reinforcement and the anticipation of correct alternatives. *Journal of Experimental Psychology* 53:417–424.

Brand, H., Woods, P. J., and Sakoda, J. M. 1956. The anticipation of reward as a function of partial reinforcement. *Journal of Experimental Psychology* 52:18–22.

Braveman, N. S., and Fischer, G. J. 1968. Instructionally induced strategy and sequential information in probability learning. *Journal of Experimental Psychology* 76:674–676.

Goodnow, J. J. 1955. Determinants of choice-distribution in two-choice situations. *American Journal of Psychology* 68:106–116.

Little, K. B., Brackbill, Y., Isaacs, R. B., and Smelkinson, N. 1963. A further test of a general utility theory model for probability learning. *Journal of Experimental Psychology* 66:107–108.

Peterson, C. R., and Ulehla, Z. J. 1965. Sequential patterns and maximizing. *Journal of Experimental Psychology* 69:1–4.

Siegel, S. 1959. Theoretical models of choice and strategy behavior: Stable state behavior in the two-choice uncertain outcome situation. *Psychometrika* 24:303–316.

SPECIFIC-PAIR INTERFERENCE ON RECALL AND ASSOCIATIVE-MATCHING RETENTION TESTS

Dennis J. Delprato
Eastern Michigan University

Retention of original-learning AB pairs was tested—by modified free recall and by associative matching—after irrelevant activity (control), unmixed AC, unmixed DC, and mixed AC/DC interpolated learning. The experimental subjects showed specific-pair interference on both tests, with no reliable difference between the mixed-list condition and the unmixed-list conditions.

One version of the interference theory of forgetting (McGovern, 1964) predicts that specific original-learning associations should be unlearned whenever identical stimuli have different responses in successive, interpolated-learning, lists—the AB-AC paradigm. But such specific associative effects have yet to be unequivocally demonstrated (Bogartz, 1965; Ceraso, 1964, Ceraso and Tendler, 1968), and continued failure to obtain them may well demand a revision of interference theory. For example, Postman has recently offered a view of retroactive inhibition, similar to an earlier position (Newton and Wickens, 1956), which emphasizes generalized interlist response competition, rather than individual stimulus-response interference (Postman and Stark, 1969; Postman, Stark, and Fraser, 1968).

Melton (1961) has outlined an attack on the question of specific associative unlearning as follows: (*a*) devise an interpolated-learning *mixed list* on which some original-learning pairs have corresponding pairs while others do not (an AB-AC/DC list, which in effect combines the AB-AC and AB-DC paradigms); then (*b*) following that interpolated learning, performance on a test of original-learning retention should be superior for the AB items in the AB-DC paradigm while the AB items in the AB-AC paradigm hould show the effects of specific associative unlearning. The rationale of this mixed-list design is that any general interlist effects (as opposed to specific interlist, or stimulus-response, effects) should influence the AB pairs of the two paradigms equally.

In the present investigation, Melton's outline was followed in order to determine whether specific-pair interference could be shown in a typical retroactive-inhibition situation. Basically, McGovern's analysis of unlearning (1964) provided the framework for two experiments.

In Experiment I, an unpaced *modified free-recall test* of original learning was used. This test, which requires recall of responses, permits measurement of the recallability of the original-learning responses of specific pairs. Previous experiments in which specific item effects were not obtained have all required some form of response recall (Bogartz, 1965; Ceraso, 1964; see also Ceraso and Tendler, 1968). It does seem possible that general interlist effects conceal specific effects on such tests; Postman's notion of functional original-learning response suppression suggests so. Indeed, we must note that absence of specific effects on such tests cannot provide conclusive evidence against specific-pair interference. For, in terms of McGovern's distinction between unlearning in the response-recall and associative phases of paired-associate learning, equivalent response recallability for DC and AC pairs precludes the manifestation of differences in associative strength when the subject is required to recall responses. Still, if tests of response recall do mask specific item interference, previous failures to obtain specific effects could be accounted for on this basis.

Experiment II used an unpaced *associative-matching test* of retention of original learning. Since all original-learning stimulus and response terms are made available to the subjects on this test, any potential response-suppression or related response-recallability influences should be eliminated. Thus, the test allows measurement of specific associative retention independent of response recallability. Whether specific-pair interference is not a factor in retroactive inhibition, or is masked on tests of response recall, should be reflected on the associative-matching test.

In addition to mixed-list conditions, typical unmixed-list groups were tested in both experiments. In an interpolated-learning *unmixed list*, there are only AC or only DC pairs. Because of this, specific and general interlist effects cannot be distinguished. But unmixed lists can be compared with mixed lists on degree of interference when the same materials were used in the two types of lists. If retention of AB pairs is not different with a mixed list (AB-AC/DC) but is different on typical unmixed lists (AB-AC and AB-DC), then only general interlist interference effects would appear to be implicated in retroactive inhibition.

METHOD

—Design—Since Experiments I and II differed only with respect to the retention test, they may be described together. In each experiment, three *experimental groups* had original learning followed by either unmixed AC (all interpolated-learning pairs AC), unmixed DC (all interpolated-learning pairs DC), or mixed AC/DC (half of the interpolated learning pairs AC, the other half DC). The retroactive-inhibition *control group* had only original learning.[1]

—Materials and notation—The eight-item paired-associate lists contained two-syllable adjective response terms that had minimal similarity of meaning. One set of eight words was used for original learning, and the other eight words were used in interpolated learning. All responses began with a different letter. Two stimulus-response pairings of the original-learning list (AB) were used equally often in each group. Stimulus terms of the original-learning list were nonsense syllables with a mean Archer (1960) association value of 62.2. Intralist stimulus similarity was kept minimal by using 16 different consonants and by using no vowel more than twice.

Original-learning AB pairs were defined by the pairs of the interpolated-learning lists. First, for each stimulus-response pairing of the AB list, unmixed-AC and unmixed-DC lists were constructed. In the DC list, three-digit numbers were used as stimuli. Intralist similarity of the numbers was kept at a minimum. The AB pairs are referred to as AC or DC pairs, depending on the kind of unmixed list received. Then, for the mixed-list condition, two versions of the interpolated-learning list were formed for each of the two versions of the AB list in such a way that each AB pair was an AC and DC item equally often. Four randomly selected items from each of the unmixed-AC and unmixed-DC lists formed one mixed list, and the remaining four items from each unmixed list served as the other mixed list for each of the two versions of the AB list. Thus, in the mixed-list condition the original-learning AB pairs that had corresponding interpolated-learning AC pairs are referred to as AC pairs; the original-learning items to which no interpolated-learning pair corresponded are referred to as DC pairs.

—Procedure—Lists were presented at a 2:2-sec rate with a 4-sec intertrial interval. There were four random orders of the pairs in each list, and each was used as the starting order equally often. All subjects learned the original list to a criterion of one perfect trial. Experimental subjects practiced on the interpolated list for 14 anticipation trials after an initial study trial. The interlist interval was 1.5 min. Control subjects spent the period of time required for interpolated learning on a digit-cancellation task.

The subjects took the original-learning retention test immediately after the digit-cancellation task or interpolated learning. The unpaced recall test of

Experiment I consisted of a booklet in which the eight stimuli of original learning were presented, one per page. The subject went through the booklet and wrote as many responses under the appropriate syllables as he could. After he indicated he was finished, the experimenter asked him to write on the back of the last page any responses from original learning that he had not already written but still could recall. The associative-matching test of Experiment II was identical to the recall test except that the original-learning responses were made available on a sheet of paper given to each subject.

—Subjects—In each experiment, 64 male and female introductory psychology students were assigned sequentially to groups of 16 subjects via block randomization.

RESULTS

The mean number of trials to the original-learning criterion in Experiment I ranged from 13.31 to 14.69 (*ps* $> .50$). In Experiment II, these means ranged from 11.06 to 14.63 (*ps* $> .10$). Thus, the groups within each experiment were adequately matched with respect to original-learning performance.

Experiment I

The mean retention scores for original learning on the recall test of Experiment I are shown in Table 1. The protocols were scored stringently (response considered correct only if paired with appropriate stimulus) and leniently (total response recall).[2]

The basic analysis of the mixed-list condition was an analysis of variance by group (experimental versus control) and pairs (AC versus DC), with the former variable a between-subjects effect and the latter a within-subjects effect. In Experiment I, the effect of groups was significant ($F = 13.14$, $df = 1/30$, $p < .005$), indicat-

Table 1. Mean retention scores on the response-recall test, Experiment I

	Experimental groups		Control group		
	AC	DC	AC		DC
Stringent scores					
Mixed list	2.44	3.25	3.69		3.81
Unmixed list	4.44	7.31		7.50	
Lenient scores					
Mixed list	2.63	3.38	3.81		3.81
Unmixed list	4.88	7.44		7.62	

ing that retroactive inhibition was obtained. The effect of pairs was also significant ($F = 5.87$, $df = 1/30$, $p < .025$), while the interaction of groups and pairs fell short of significance ($F = 3.15$, $df = 1/30$, $p > .05$).

Individual comparisons are of critical importance to questions about specific associative effects in the mixed list.[3] As suggested in the introduction, if specific-pair interference is a factor in retroactive inhibition, recall of AC pairs should be inferior to recall of DC pairs in the mixed-list condition. Specific-pair interference was found, in that these subjects recalled significantly fewer responses from AC pairs than from DC pairs ($F = 8.80$, $df = 1/30$, $p < .01$). The AC versus DC comparison within the control group was far from significant ($F < 1$). Other individual comparisons in the mixed-list condition showed reliable differences between the experimental and control groups (retroactive inhibition) for both AC and DC pairs ($F = 22.30$, $df = 1/30$, $p < .001$, and $F = 4.52$, $df = 1/30$, $p < .05$, for AC and DC respectively).

Significant variation was found among the means in the unmixed-list conditions ($F = 24.88$, $df = 2/45$, $p < .001$). Newman-Keuls individual comparisons indicated significant retroactive inhibition for AC ($p < .01$) but not for DC ($p > .05$). As is typically the case with unmixed lists, AC was inferior to DC ($p < .01$).

In order to compare the amount of interference in the mixed list with that in the unmixed list, the mixed-list scores were converted to a base of 8, then mixed list versus unmixed list differences were analyzed for AC and DC separately. This transformation yielded means of 4.88 and 6.50 for mixed-list AC and DC pairs respectively. Amount recalled was not found to be reliably different in the two types of list (ps $> .10$), which suggests that specific-pair interference was also evident on the unmixed lists, since (*a*) specific effects were obtained on the mixed lists and (*b*) equivalent recall scores were obtained on mixed and unmixed lists.

Experiment II

The mean numbers of original-learning pairs correctly matched in Experiment II are presented in Table 2. Analysis of the mixed-list condition yielded a significant effect of group (experimental versus control; $F = 16.20$, $df = 1/30$, $p < .001$), of pairs (AC versus DC; $F = 8.59$, $df = 1/30$, $p < .01$), and of their interaction ($F = 8.59$, $df = 1/30$, $p < .01$). Specific-pair effects were obtained on the associative-matching test, since these subjects matched significantly fewer AC pairs than DC pairs ($F = 17.18$, $df = 1/30$, $p < .001$). While reliable retroactive inhibition was indicated for AC pairs of the mixed list ($F = 24.41$, $df = 1/30$, $p < .001$), it was not found on DC pairs ($F = 3.03$, $df = 1/30$, $p > .05$).

Table 2. Mean retention scores on the associative-matching test, Experiment II

	Experimental groups		Control group	
	AC	DC	AC	DC
Mixed list	2.81	3.50	3.88	3.88
Unmixed list	6.19	7.19	7.75	

The results for the unmixed-list conditions were similar to those of the mixed-list group (and to the unmixed-list groups of Experiment I). That is, significant retroactive inhibition was found with AC interpolation ($p < .01$), but not with DC interpolation ($p > .05$). Also, original-learning matching was significantly better in the DC than in the AC group ($p < .05$).

When the mixed-list scores were converted to a base of 8, the means were 5.63 and 7.00 for the AC and DC pairs respectively. These transformed scores of the mixed-list condition were not significantly different from the corresponding scores of the unmixed-list groups ($ps > .30$). Thus, as in the recall test of Experiment I, proportion of interference did not differ for mixed and unmixed lists. If specific-pair interference accounts for the associative-matching differences on the mixed list, it seems appropriate to conclude that specific-pair interference was also evident with the unmixed lists.

DISCUSSION

Assuming that Melton's outline is sound, the data of both Experiments I and II support the notion that interference between specific interlist pairs is a factor in retroactive inhibition. The speculation that specific-pair interference is masked on tests of response recall was not substantiated, since such interference did influence response recallability in Experiment I. In addition, when response-recallability effects were eliminated on the associative-matching test of Experiment II, specific interference was obtained. It is important to note that the present findings in no way deny that general interlist effects are factors in retroactive inhibition; they do indicate that recent attempts to minimize specific associative unlearning (Postman and Stark, 1969; Postman, Stark, and Fraser, 1968) may be premature.

A certain theoretical difficulty within interference theory has justified searches for alternatives to specific associative unlearning. There has been a close relationship between the hypothesis of specific-pair unlearning and the elicitation-extinction hypothesis (Keppel, 1968; McGovern, 1964). The latter has been considered the mechanism of unlearning and assumed to be both association-specific, resulting in stimulus-response unlearning, and general (context-response), resulting in response unlearning—and the results of several recent experiments have demonstrated that it has serious shortcomings (Birnbaum, 1968; Tulving and Madigan, 1970; Zavortink and Keppel, 1969). However, it should be pointed out that the apparent failure of the elicitation-extinction hypothesis does not imply that specific associative effects are unimportant sources of interference. And although generalized interlist response competition can perhaps account for losses in response recallability, an alternative to the elicitation hypothesis as a mechanism for specific associative unlearning is still needed.

One alternative is that the requirement to learn new response C to original-learning stimulus A is conducive to loss of implicit mediators that underlie AB associative hookup. Although Adams and Montague (1967) found no decrease in number of reported original-learning AB mediators after AC practice (in direct contrast to this alternative hypothesis), they used words as stimulus and response terms in their lists, and it may well be that mediators between such highly meaningful items are more resistant to retroactive loss than are mediators between items of lower meaningfulness.

The suggestion that loss of mediators is critical for specific associative unlearning has several testable implications. First, obviously, AB-AC practice should adversely affect retention of original-learning mediators. Second, the amount of specific associative unlearning should vary inversely with number of original-learning mediators available at the time of retention, and the relationship between available mediators and retention must be pair-specific. Third, if mediators are more resistant to disruption as meaningfulness of the members of original-learning pairs increases, specific associative unlearning should vary inversely with stimulus and response meaningfulness. Indirect evidence that mediational activity is involved in the resistance of specific pairs to unlearning has been obtained by M. H. Garskof (1968) and Garskof, Sandak, and Malinowski (1965), who found that instructions to mediate decreased stimulus-response unlearning on an associative-matching retention test. But whether or not specific-pair interference is a matter of the loss of mediator availability, it seems clear that specific associations are liable to unlearning effects.

Notes

The research was supported in part by Grant HD02626-02 from the National Institutes of Health to B. E. Garskof. Portions of Experiment II were presented at the meetings of the American Psychological Association, Washington, D. C., September, 1969. Received for publication October 1, 1970.

1. Since construction of the interpolated-learning list distinguished between mixed- and unmixed-list conditions, a single group served as the (retroactive-inhibition) control group for both conditions.

2. Only the results of the analysis of stringent scores are discussed, since the analysis of lenient scores led to essentially the same conclusions.

3. These comparisons followed Winer (1962, pp. 310–311).

References

Adams, J. A., and Montague, W. E. 1967. Retroactive inhibition and natural language mediation. *Journal of Verbal Learning and Verbal Behavior* 6:528–535.

Archer, E. J. 1960. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs* 74(10, whole no. 497).

Birnbaum, I. M. 1968. List 1 recall and associative strength of List 2 pairs. *American Journal of Psychology* 81:442–447.

Bogartz, W. 1965. Unlearning of specific associations in retroactive inhibition. *Proceedings of the 73rd Annual Convention of the American Psychological Association* 53–54.

Ceraso, J. 1964. Specific interference in retroactive inhibition. *Journal of Psychology* 58:65–77.

Ceraso, J., and Tendler, M. 1968. Pair vs. list interference. *American Journal of Psychology* 81:47–52.

Garskof, B. E., Sandak, J. M., and Malinowski, E. 1965. Controlling the "fate" of first list associates. *Psychonomic Science* 2:315–316.

Garskof, M. H. 1968. Short-term retention of paired associates as a function of instructions and retention measure. *Journal of Verbal Learning and Verbal Behavior* 7:409–412.

Keppel, G. 1968. Retroactive and proactive inhibition. In *Verbal behavior and general behavior theory,* ed. T. R. Dixon and D. L. Horton, pp. 172–213. Englewood Cliffs, N. J.: Prentice-Hall.

McGovern, J. B. 1964. Extinction of associations in four transfer designs. *Psychological Monographs* 78(16, whole no. 593).

Melton, A. W. 1961. Comments on Professor Postman's paper. In *Verbal learning and verbal behavior,* ed. C. N. Cofer, pp. 179–193. New York: McGraw-Hill.

Newton, J. M., and Wickens, D. D. 1956. Retroactive inhibition as a function of the temporal position of the interpolated learning. *Journal of Experimental Psychology* 51:149–154.

Postman, L., and Stark, K. 1969. Role of response availability in transfer and interference. *Journal of Experimental Psychology* 79:168–177.

Postman, L., Stark, K., and Fraser, J. 1968. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior* 7:672–694.

Tulving, E., and Madigan, S. A. 1970. Memory and verbal learning. *Annual Review of Psychology* 21:437–484.

Zavortink, B., and Keppel, G. 1969. Retroactive inhibition as a function of list 2 study and test intervals. *Journal of Experimental Psychology* 81: 185–190.

THE EFFECTS OF EXPECTED VALUE DIFFERENCE AND EXPECTED REGRET RATIO ON PREFERENCE STRENGTH

Wayne Lee
Bell Telephone Laboratories, Holmdel, New Jersey

In two experiments on the effects of expected value difference (D) and expected regret ratio (R) on the strength of preference between gambles, D and R were unconfounded. It was found that D can affect preference strength when R is controlled, and vice versa. The results, however, show that there are strong determinants of preference strength other than D and R—determinants as yet not satisfactorily accounted for.

In another paper (1971), Lee has discussed two indices that have been related to strength of preference for one gamble over another. In particular, the paper concerned choice situations that can be characterized by the payoff matrix:

$$\begin{array}{c|c|c|} & P & 1-P \\ \hline \alpha & a & b \\ \hline \beta & c & d \\ \hline \end{array}$$

On each trial, the subjtct must choose one of two alternatives, α or β (α and β are also called gambles). There are two so-called states of nature, corresponding to the two columns. The left column obtains with probability P, and the right column obtains with probability $1 - P$. The cells a, b, c, and d are the possible payoffs to the subject (usually real or make-believe monetary amounts). The particular payoff he receives on a given trial depends on his choice and on the state of nature, which is not revealed until after he makes his choice.

A number of experimental studies have been concerned with what might be called the subject's *strength of preference* for one possible choice over the other (Edwards, 1956; Mosteller and Nogee, 1951; Myers, Reilly, and Taub, 1961; Myers, Suydam, and Gambino, 1965; Suydam, 1965; Taub and Myers, 1961). The strength of preference has been measured in two ways: by the probability, p $(\alpha\,\beta)$, that α is chosen rather than β; and by the subject's direct ratings of his strength of preference for α over β (or vice versa).

Two indices have been proposed to account for the different preference strengths found as a function of variation in payoffs and state-of-nature probabilities. One is the *difference in expected value* (D) for the two gambles. The other index is the *expected regret ratio* (R), which has also been called "expected loss ratio" and "relative expected loss" (Suydam, 1965; Edwards, 1956). Let $x = a - c$ and $y = d - b$. When x and y are positive, the expected value difference, $EV_\alpha - EV_\beta$, is

$$D(\alpha\beta) = Px - (1 - P)y \qquad [1]$$

and expected regret ratio is

$$R(\alpha\beta) = \frac{Px}{Px + (1 - P)y} \qquad [2]$$

(for a more detailed discussion of these indices, see Lee, 1971).

Previous research has shown that preference strength increases with D (Mosteller and Nogee, 1951; Myers, Reilly, and Taub, 1961; Suydam, 1965; Myers, Suydam, and Gambino, 1965) and with R (Edwards, 1956; Taub and Myers, 1961; Suydam, 1965; Myers, Suydam, and Gambino, 1965). However, in previous research D and R have been strongly confounded, and the independent effects of these two indices have not been determined. Lee (1971) showed how to construct sets of payoff matrices orthogonal (unconfounded) in D and R. The present paper reports two experiments designed to assess the independent effects of D and R on preference strength.

EXPERIMENT I

Method

A basic set of nine payoff matrices orthogonal in D and R was constructed, using three levels of $D(\alpha\beta)$ (4, 8, and 16), each combined with three levels of $R(\alpha\beta)$ (.6, .75, and .9).[1] P was 1/2. As Lee has shown, for given values of $D(\alpha\beta)$, $R(\alpha\beta)$, and P, one can compute

$$x = \frac{D(\alpha\beta)R(\alpha\beta)}{P[2R(\alpha\beta) - 1]} \quad \text{and} \quad y = \frac{D(\alpha\beta)[1 - R(\alpha\beta)]}{(1 - P)[2R(\alpha\beta) - 1]}.$$

Since $x = a - c$ and $y = d - b$, further constraints are necessary to determine the payoffs. For the basic set (set A) of matrices, b and c were arbitrarily set equal to 0. Two other sets of nine gamble pairs were formed by transforming the first set. Set B was formed by adding 10 to each payoff in set A, and set C was formed by adding -10 to each payoff in set A. Such transformations do not affect x and y, so they leave the D and R values unchanged. Also, since the same constant was added to each column of the payoff matrix, the variance of each gamble (α and β) remained unchanged under the transformations. The 27 payoff matrices thus derived are shown in Table 1.

Each subject had to choose once between α and β for each of the 27 payoff matrices, and to rate his strength of preference for each choice over the other alternative. Table 2 illustrates the method of presentation for each gamble pair. Each subject received a booklet with 27 pages, with one such presentation on a page. A different random order of presentation was used with each booklet.

The subject was told that the state of nature for each payoff matrix would be determined by throwing a die with WUH printed on three sides and ZEQ printed on the other three sides. Such a die has been shown to yield subjective probabilities of 1/2 for each possible outcome (Davidson, Suppes, and

Table 1. Payoffs for gamble pairs, Experiment I

		R					
	D	.6		.75		.9	
Set A							
	4	24	0	12	0	9	0
		0	16	0	4	0	1
	8	48	0	24	0	18	0
		0	32	0	8	0	2
	16	96	0	48	0	36	0
		0	64	0	16	0	4
Set B							
	4	34	10	22	10	19	10
		10	26	10	14	10	11
	8	58	10	34	10	28	10
		10	42	10	18	10	12
	16	106	10	58	10	46	10
		10	74	10	26	10	14
Set C							
	4	14	−10	2	−10	− 1	−10
		−10	6	−10	− 6	−10	− 9
	8	38	−10	14	−10	8	−10
		−10	22	−10	− 2	−10	− 8
	16	86	−10	38	−10	26	−10
		−10	54	−10	6	−10	− 6

Table 2. Presentation of a gamble pair

WUH	ZEQ		
			______very strong preference
			______definite preference
24	0	A	______moderate preference
			______slight preference
			______slight preference
			______moderate preference
0	16	B	______definite preference
			______very strong preference

Siegel, 1957). No such die was actually present. The subjects were told that the die would be thrown 27 times afterwards and that the payoff for each matrix would be determined jointly by their choices and by the outcome from the die throw.

The subjects were told to imagine that the payoffs were amounts of money, though they realized that no money would actually change hands. Twenty-three of them were told that the payoff matrix entries were pennies, and the rest were told they were dollars. In essence, payoffs for the latter group were obtained from those of the former group by use of a constant multiplier, 100. Such a constant multiplier does not affect the *R* of a payoff matrix; *D*, however, is altered by the same ratio as the individual payoffs.

The gamble with the higher expected value was not always placed in the top row of the matrix. Its placement was random across subjects and payoff matrices. The top row, however, was always labeled A. The subject had to choose A or B. No indifferent response was allowed.

Twenty-two subjects were run under the 'dollars' instructions, and 23 were run under the 'pennies' instructions. They were run in subgroups averaging about five, and all in a group received the same instructions. Both male and female college students were utilized. They were sent by the University of California employment center to participate for a fixed stipend.

Results

Five subjects in each instructional group were eliminated from the analysis below because they gave only 'very strong preference' for the higher expected-value choice. Although such responses are sensible, they do not contribute to *D*, *R*, or set effects. Three more subjects were eliminated from the dollars group: one chose A (the upper row) every time; one appeared to choose completely randomly, and one did a lot of erasing, against instructions. This left 14 subjects in the dollars group and 18 in the pennies group for the analysis of variance.

Each response was coded as an integer from 1 through 8: 8 meant that the subject had a 'very strong preference' for the higher expected-value choice, and 1 meant he had a 'very strong preference' for the lower expected-value choice. Intermediate responses were assigned values of 2 through 7, with the higher scores indicating greater degrees of attraction to the higher expected-value gamble.

An analysis of variance was performed on these scores. Factors in the analysis were *D* (three levels), *R* (three levels), set of nine bet pairs (three levels), and instructions (two levels). Only instructions was a between-subjects factor.

The only significant effects were the main effects for *D* [$F(2/60) = 10.3$; $p < .005$] and set [$F(2/60) = 7.2$; $p < .005$]. The mean ratings for the three levels of *D*, from low to high, were 6.81, 6.94, and 7.24. The greater the difference in expected value, the more the higher expected-value gamble was preferred to the other. The mean ratings were linear with expected value difference. The mean ratings for sets A, B, and C were 7.16, 6.99, and 6.82 respectively. The basic set (with zeros along a diagonal) had the highest average rating. Perhaps preferences were more clear for these pairs, since it was obviously necessary to compare only two numbers to decide on preference. It also appears that subjects give higher preference ratings when they cannot lose (set B, which has only positive entries in the matrix) than when losses are probable (set C).

What happened to the *R* effect that one would expect from previous studies? Though the effect of this factor did not reach conventional levels for acceptance, the mean ratings for the three levels of *R* (.6, .75, and .9) were 6.91, 7.01, and 7.06 respectively, which is the expected order.

EXPERIMENT II

The lack of a significant *R* effect in Experiment I did not appear to be in accord with previous experimental results. Since an *R* effect was suggested by the data, however, it seemed possible that an *R* effect could be obtained if a greater range for *R* were employed. Such an increment in range would not result in more excessive *R* values than previous experimenters have employed. Therefore, it was decided to perform a second experiment similar to the first but using a larger range for *R*. Some other features were changed too.

Method

A second set of gamble pairs was constructed, with *R* values of .55, .75, and .95. Also, an attempt was made to reduce the number of subjects who would always have a very strong preference for the higher expected-value gamble

Table 3. Payoffs for gamble pairs, Experiment II

		R					
	D	.55		.75		.95	
Set A′							
	.5	5.50	.00	1.50	.00	1.06	.00
		.00	4.50	.00	.50	.00	.06
	1.0	11.00	.00	3.00	.00	2.11	.00
		.00	9.00	.00	1.00	.00	.11
	1.5	16.50	.00	4.50	.00	3.17	.00
		.00	13.50	.00	1.50	.00	.17
Set B′							
	.5	3.50	2.00	− .50	2.00	− .94	2.00
		−2.00	6.50	−2.00	2.50	−2.00	2.06
	1.0	9.00	2.00	1.00	2.00	.11	2.00
		−2.00	11.00	−2.00	3.00	−2.00	2.11
	1.5	14.50	2.00	2.50	2.00	1.17	2.00
		−2.00	15.50	−2.00	3.50	−2.00	2.17
Set C′							
	.5	7.50	−2.00	3.50	−2.00	3.06	−2.00
		2.00	2.50	2.00	−1.50	2.00	−1.94
	1.0	13.00	−2.00	5.00	−2.00	4.11	−2.00
		2.00	7.00	2.00	−1.00	2.00	−1.89
	1.5	18.50	−2.00	6.50	−2.00	5.17	−2.00
		2.00	11.50	2.00	− .50	2.00	−1.83

by utilizing smaller D values than before, namely, .5, 1.0, and 1.5. As before, a basic set (set A′) of gamble pairs was constructed from these specifications by setting $b = c = 0$.

As before, two alternate sets were constructed from the basic set. Set B′ was constructed by adding −2.00 to the first column, and +2.00 to the second column of each matrix in the basic set. Such transformations affect neither D nor R. Unlike the alterations of Experiment I, these transformations left the average payoff in a matrix unchanged. On the other hand, these transformations did change the variances of the gambles, whereas the transformations in Experiment I did not. Set C′ was formed by adding +2.00 to the first column and −2.00 to the second column. Note that in set B′ the higher expected-value gamble has the lower variance, but in set C′ it has the higher variance. Sets A′, B′, and C′ are shown in Table 3.

The procedure was the same as in Experiment I, except for these differences. In contrast to Experiment I, the payoff matrix entries were not integers; therefore, values in the payoff matrix were given to two decimal places. The subjects in the 'dollars' group were told that each entry was in dollars and cents; those in the 'pennies' group were told that each entry was in pennies, with a fraction of a penny given to the right of the decimal point, and that all such fractional parts would be included in summing the final score.

There were 20 subjects in each group. Four were eliminated from the analysis: two because they gave only 'very strong preference' for the higher expected-value gamble, and two because they did not follow instructions. This left 36 subjects (18 in each group) on which the analysis of variance was performed.

Results

As in Experiment I, each subject had a score from 1 to 8 for each gamble pair. The design was the same as in Experiment I: D (three levels), R (three levels), set (three levels), and instructions (two levels). Only instructions was a between-subjects variable.

Three main effects were significant: those of D $[F(2/68) = 9.5, p < .005]$, R $[F(2/68) = 82.0, p < .005]$, and set $[F(2/68) = 74.4, p < .005]$. The following two-factor interactions were significant: those of D and set $[F(4/136) = 7.19, p < .005]$ and of R and set $[F(4/136) = 26.40, p < .005]$. Also, the three-factor interaction of D and R and set was nearly significant at the .05 level.

Table 4 gives the mean ratings for the levels of D, R, and set. Surprisingly, the mean rating for $D = 1.5$ was less than the mean rating for $D = 1.0$. The mean ratings for the R levels differed appreciably, and in the expected order. The effect of set was strongest of the three. The highest ratings were for the basic set with the zero diagonals. The next highest ratings were for set B′, in which the higher expected-value gambles had lower variances. The lowest ratings were for set C′, in which the higher expected-value gambles had higher variances.

Table 5 gives mean ratings as functions of D, R, and set. Each entry gives the average rating across all subjects in the analysis (for both instructional groups). The values yielding the interaction of D and set are the nine means at the right of the table. Note that D had essentially no effect in set A′, had a monotonic effect in set B′, and a nonmonotonic effect in set C′. This nonmonotonicity was consistent for the three R levels.

The values yielding the interaction of R and set are given by the three rows of means. Note that mean ratings were monotonic with

Table 4. Mean ratings for the levels of *D*, *R*, and set, Experiment II

D	Rating	*R*	Rating	Set	Rating
.5	6.44	.55	5.92	A′	7.51
1.0	6.78	.75	6.76	B′	6.74
1.5	6.66	.95	7.20	C′	5.62

Table 5. Mean ratings as a function of *D*, *R*, and set, Experiment II

		R			
	D	.55	.75	.95	Mean
Set A′					
	.5	7.42	7.39	7.70	7.50
	1.0	7.33	7.45	7.72	7.50
	1.5	7.39	7.61	7.61	7.54
	Mean	7.38	7.48	7.68	
Set B′					
	.5	5.50	6.83	6.86	6.40
	1.0	6.33	6.97	7.03	6.78
	1.5	6.95	7.03	7.20	7.06
	Mean	6.26	6.94	7.03	
Set C′					
	.5	4.00	5.61	6.67	5.43
	1.0	4.58	6.45	7.11	6.05
	1.5	3.75	5.50	6.92	5.39
	Mean	4.11	5.85	6.90	

$R(\alpha\beta)$ for all three sets. The effect of *R* was relatively modest for set A′, but all the mean ratings were high (the maximum possible was 8.00), so there was little room for differentiation. In set B′ the effect of *R* was strong, although the second and third levels of *R* hardly differed. The effect of *R* was strongest in set C′, and large differences appeared between all three *R* levels.

The extreme strength of the effect of *R* in set C′ was largely due to the surprisingly low ratings for $R = .55$; in fact, each of the three gamble pairs involved had ratings indicating indifference between gambles (indifference is at 4.5) or preference for the smaller expected-value gamble. Examination of these three gamble pairs (Table 3) shows that there is exactly one negative payoff for each pair, and it is associated with the higher expected-value gamble. No other gamble pair in the experiment had this characteristic, and no other gamble pair had a mean rating less than 5.00.

No such structural feature can explain the nonmonotonicity for *D* in set C′. The structure (in terms of negative and positive entries) for each gamble pair for $D = 1.5$ is the same as the gamble pair above it (for $D = 1.0$).

DISCUSSION

We set out to determine the independent effects of *D* and *R*. Both *D* and *R* can affect preference strength, but the effects are so dependent on the choice of the arbitrary payoff parameters b and c that no satisfactory account of mean ratings can be obtained on the basis of D and R alone. It is peculiar, then, that R has been found in previous research to predict preference strength so well (Edwards, 1956; Taub and Myers, 1961; Suydam, 1965; Myers, Suydam, and Gambino, 1965). In part this was due to the scoring methods used[2] and to the confounding of R with D. Additional explanation seems necessary, however. In previous research, each gamble had one possible positive payoff and one possible negative payoff. Simpler results may occur with such gambles than with the less restricted gambles used in the present experiments. Furthermore, previous authors utilized a variety of P values, whereas P was always 1/2 in the present experiments. Suydam (1965) reported that the only notable failures for R as a predictor in the data of Myers, Reilly, and Taub (1961) were for $P = 1/2$.

The divergent results between Experiments I and II can be accounted for, to a large extent, by the different manner in which the alternative sets were constructed. However, both experiments had a set with zero diagonals, and one would hope for consistency here. Why then, for set A′ (Experiment II) was there no effect of D (see Table 5), yet in Experiment I a main effect for D but not an interaction, implying that a D effect existed for set A? Examination in detail of the data of Experiment I showed that the effect of D was weakest for set A: the range for the mean ratings was only .22, versus .40 for set B and .64 for set C. Thus, even though these differences did not produce a significant interaction, they do help reconcile the two experiments: D had little or no effect for the sets with zero diagonals in both experiments.

Similarly, there was a significant R effect in Experiment II but not in Experiment I. There was, of course, a larger R range in Experiment II, which was put in to bring out an effect. This greater range may have been causative. However, the difference in the effect of R between the experiments was less for sets A and A′ than for all sets combined. The effect of R in A′ was .30 (between the lowest and highest R levels), and in set A it was .24. Thus, it may not have been primarily (if at all) the increase in R range that produced the large R effect in Experiment II but, rather, the particular forms for the gamble pairs in sets B′ and C′.

It may seem surprising that the ratings did not differ between instructional groups in either experiment, since the D range for the dollars group was 100 times the D range for the pennies group. This undoubtedly resulted because of well-known tendencies of subjects to spread their categorical judgments across the range provided rather than interpreting a scale in an absolute sense. Still, if there were a nonlinear utility function affecting preference strengths, some differences between the instructional groups might well have appeared. This suggests that postulating a nonlinear utility function will not be helpful in explaining the results found.

It might be thought that variance preferences (Combs and Pruitt, 1960; Lichtenstein, 1965) could explain some of the results of the present experiments. Perhaps so. But there were no variance alterations between sets in Experiment I, so the main effect for set in Experiment I cannot be attributed to variance preferences.

To sum it up, these two experiments were run to assess the effects of expected value difference (D) and expected regret ratio (R) on strength of preference between gambles. In contrast to previous investigations, D and R were unconfounded; and D was found to affect preference strength when R was controlled, and vice versa. Whereas previous investigations suggested that R, or R with D, was adequate to predict relative preference strength with high accuracy, the present results, using less restricted types of gambles, demonstrate that there are strong determinants of preference strength other than D and R. A satisfactory accounting of these determinants remains to be accomplished.

Notes

The author thanks Rohn Schoss for his assistance in running the experiments and analyzing the data. The author's present address is 1393E Solano Avenue, Albany, California 94706. Received for publication June 22, 1970.

1. For an orthogonal set of payoff matrices, the $D(\alpha\beta)$s must all be positive or must all be negative. If they are positive, the $R(\alpha\beta)$s must be between 1/2 and 1 (see Lee, 1971).

2. In calculating the indices D and R, and measuring the preference strength for a pair of gambles, the investigator, to put it simply, must decide which gamble is α and which is β. Previous authors made these decisions in an arbitrary manner. The author explains (Lee, 1971) why it is preferable to let α always be the gamble with the higher expected value.

References

Coombs, C. H., and Pruitt, D. G. 1960. Components of risk in decision making: Probability and variance preferences. *Journal of Experimental Psychology* 60:265–277.

Davidson, D., Suppes, P., and Siegel, S. 1957. *Decision making.* Stanford, Calif.: Stanford University Press.

Edwards, W. 1956. Reward probability, amount and information as determiners of sequential two-alternative decisions. *Journal of Experimental Psychology* 52:177–188.

Lee, W. 1971. Preference strength, expected value difference, and expected regret ratio. *Psychological Bulletin* 75:186–191.

Lichtenstein, S. 1965. Bases for preferences among three-outcome bets. *Journal of Experimental Psychology* 69:162–169.

Mosteller, F., and Nogee, P. 1951. An experimental measurement of utility. *Journal of Political Economy* 59:371–404.

Myers, J. L., Reilly, R. E., and Taub, H. A. 1961. Differential cost, gain and relative frequency of reward in a sequential choice situation. *Journal of Experimental Psychology* 62:357–360.

Myers, J. L., Suydam, M. M., and Gambino, B. 1965. Contingent gains and losses in a risk-taking situation. *Journal of Mathematical Psychology* 2:363–370.

Suydam, M. M. 1965. Effects of cost and gain ratios, and probability of outcome on ratings of alternative choices. *Journal of Mathematical Psychology* 2:171–179.

Taub, H. A., and Myers, J. L. 1961. Differential monetary gains in a two-choice situation. *Journal of Experimental Psychology* 61:157–162.

UNIVARIATE AND BIVARIATE INFORMATION STORAGE

Stefan Slak
University of Toledo

For 120 undergraduates, average memory span in bits of information was higher for bivariate figures (color and shape both varying; 16 alternatives) than for univariate figures (color or shape varying; 4 alternatives). But the gain in information storage was inferior to the gains obtained by combining two dimensions in information-transmission experiments.

The channel capacity for multidimensional absolute judgment is known to be higher than the channel capacity for unidimensional judgment, though it falls short of the sum of the respective channel capacities for individual dimensions (Miller, 1956; Slak, 1969). The question arises whether a similar relationship holds for short-term information storage capacity. To answer this question, it is necessary to devise items that can be varied in one or more than one dimension. At least two sets of univariate items are needed, each set variable in one of the two dimensions; and one set of bivariate items, independently variable in both dimensions. The purpose of this study was to compare the univariate and bivariate information storage, as measured by memory span for figural stimuli expressed in bits of information.

METHOD

—Subjects—Introductory psychology students, 60 males and 60 females, served as subjects, as part of the course requirement.

—Materials—There were 135 lists of figural items. The figures were drawn on tapes for presentation on a Stowe memory drum. The dimensions in which figures could vary were shape (square, triangle, circle, or hexagon) and color (black, green, blue, or red). Each item of a list was a figure with a given shape and in a given color. Only contours of figures were drawn. The size of the figures was such that each occupied an approximate area of 1/2 by 1/2 in.

In one-third of the lists, only shape varied, while the color remained black. In another third of the lists, only color varied, while shape remained constant (a circle). The remaining lists consisted of bivariate figures, independently varying in both shape and color. The order of items in a list was random.

Lists were of nine different lengths, from 2 to 10 items in a list. Five different lists were made for each condition (figure-variable, color-variable, and bivariate) and each length of the list. A supply of blank index cards and color pens (red, black, green, and blue) was provided for the subjects' use.

—Procedure—Three different types of lists corresponded to the three different conditions: the figure-variable and the color-variable (both univariate) conditions, and the bivariate condition. Twenty subjects of each sex were assigned to each of the three experimental conditions, so that the design corresponds to 2 × 3 factorial, with sex and condition as factors.

In a given trial, a subject was presented with a list on a Stowe memory drum at a 3-sec rate. Immediately after the end of list presentation, he had 30 sec to reproduce the figures by drawing them on a blank index card in their exact serial order. In the color-variable and bivariate conditions, he had at his disposal pens of appropriate colors, positioned in a constant order. (In a few pilot experiments, the subjects were observed as they used the pens; it was concluded that manipulating the pens would not cause any major difficulties.) Only perfect reproductions of items in serial order were scored as correct. First, a subject was given five lists of length two, then five lists of length three, and so on, until he failed to reproduce correctly all five lists of a given length. For each subject, memory span was determined by adding .2 to 1.5 for every correct reproduction of a list; that is, for each list length the same proportion of the unit interval was added to 1.5 (the lower limit of the interval corresponding to length of two items) as the proportion of lists of a given length correctly reproduced.

Immediately after the experiment, the subject was questioned about the strategies he had used in memorizing the items, with particular emphasis on the implicit use of verbal labels for colors or shapes.

In univariate lists, an item is worth two bits of information, since there are 4 equiprobable alternatives. In bivariate lists, where there are 16 equiprobable alternatives, an item is worth four bits of information.

RESULTS

In order to express memory span in bits of information, memory spans in the univariate conditions were multiplied by two, those in the bivariate condition by four. Mean memory spans in bits of information for the three conditions and both sexes are presented in Table 1. Factorial analysis of variance yielded a significant effect of condition [$F(2, 114) = 36.52, p < .01$], with no significant sex difference [$F(1, 114) = 0.21$] or interaction of condition and sex [$F(2, 114) = 0.41$]. In further analysis, the two sex groups were pooled. There was

Table 1. Mean memory span in bits of information

	Figure-variable condition	Color-variable condition	Bivariate condition	All conditions
Males	12.36	12.14	15.36	13.29
Females	11.78	12.32	15.28	13.13
Both sexes	12.07	12.23	15.32	

no significant difference between the two univariate conditions $[t(78) = 0.37, p > .05]$. The most relevant comparison is the difference between univariate and bivariate conditions. Both of the former, the figure-variable condition $[t(78) = 8.91, p < .01]$ and the color-variable condition $[t(78) = 11.76, p < .01]$, reliably differed from the bivariate condition. The confidence interval at the .01 level for the mean memory span in the bivariate condition was 15.32 ± .25.

The number of subjects who used implicit verbal labels for colors or shapes was recorded. There were nine such subjects in the figure-variable, six in the color-variable, and eight in the bivariate condition. Only negligible, nonsignificant differences occurred between subjects who covertly verbalized and those who did not report covert verbalizations of items.

DISCUSSION

The results show that bivariate information storage, as measured by memory span for figural stimuli in bits of information, was superior to univariate information storage.[1] If memory span for figural stimuli were independent of the information content per item, it would be twice as high in the bivariate condition, since bivariate items contain twice as much information as do the univariate items. It was not twice as high, as Table 1 shows. Thus, the conclusion from our findings—the conclusion that bivariate short-term memory capacity for figural stimuli in terms of bits of information is superior to the univariate one but falls short of the sum of the respective univariate short-term memory capacities for the two individual dimensions—parallels the previous findings on unidimensional and multidimensional absolute judgment.

However, a closer examination reveals an apparent difference between the data on absolute judgment and the data of the present study. As Slak (1969) showed, the channel capacity for multidimensional information transmission may be expressed as a function of the unidimensional channel capacities, so that the former may be defined as the distance of a point in the multidimensional euclidean space from the origin, when the point is defined in terms of respective unidimensional channel capacities as coordinates. If the model is applied to the two-dimensional information storage, the memory span in terms of bits of information for the bivariate condition should have been equal to $(12.07^2 + 12.23^2)^{1/2} = 17.18$, which is beyond the upper limit of the confidence interval for the mean bivariate memory span that was in fact obtained. Thus, the gain in information storage for figural stimuli as a function of combining variable dimensions was less than the gain in information transmission when dimensions are combined in absolute judgment.

If A and B stand for memory spans in bits of information for the two univariate conditions, and C for memory span in bits for the bivariate condition, the relationship between univariate and bivariate information storage for figural stimuli may be expressed by the inequality

$$\tfrac{1}{2}(A + B) < C < (A^2 + B^2)^{1/2}.$$

The quantity in the left portion of the inequality corresponds to the hypothesis of a constant memory span in bits of information, and the quantity in the right portion corresponds to the bivariate memory span that would be expected if the combining of dimensions had the same effect for figural stimuli as in informational studies on absolute judgment.

Reported implicit verbalization did not seem to have any effect on performance. Because of the nature of the design and the type of data (verbalization data were only supplementary and tentative), it was impossible to ascertain to what extent the processes involved in memorizing the figural materials were purely perceptual and to what extent verbal. It is doubtful that the distinction is fully justified, however, for in normal adults, word-referent connections for the type of material used are so strongly overlearned that such figures may represent conceptual units with verbal and perceptual characteristics as component parts, whether subjects are aware of verbal 'contamination' or not.

Notes

The author thanks Julian Mingus, graduate assistant in psychology, for the data collection, and Robert Carr, graduate assistant in psychology, for preparation of the materials. Received for publication September 9, 1970.

1. It may be noted, however, that memory span in terms of the number of *figures* correctly *reproduced* was less in the bivariate condition than in either univariate condition.

References

Miller, G. A. 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63:81–97.

Slak, S. 1969. Multidimensional information transmission as a function of unidimensional channel capacity. *Psychological Review* 76:224–226.

IS AUTOKINETIC MOTION RELATED TO EXPERIENCED DISPLACEMENT?

Joseph Glick
University of Minnesota, Minneapolis

Experiment I investigated both apparent visual movement and apparent displacement of a luminous target under the conditions by which Gregory and Zangwill produced autokinetic motion. Experiment II investigated changes in egocentric localization (apparent displacement) measured by nonvisual means under the same conditions. In Experiment I the measures of motion and of displacement were indeed related. Experiment II further suggested that the relationship holds beyond the purely visual sphere.

Autokinetic motion is the experience of motion of an objectively stationary object. In some cases the phenomenal appearance of this 'illusory' motion is identical to the appearance of real motion—as a displacement in space over time. In other cases the motion is reported to be 'paradoxical'—an experience of motion without a corresponding experience of displacement (Moustgaard, 1963; Gregory and Zangwill, 1963).

The distinction between apparently real and apparently paradoxical motion has had important consequences for the theoretical treatment of the autokinetic phenomenon. Several theorists (Carr, 1910; Brosgole, 1967) have drawn analogies between autokinetic motion and real motion. They have, accordingly, attempted to explain autokinetic motion on the analogy to real motion. Since real motion is generally assessed relative to a reference system or framework, autokinetic motion is seen as reflecting perceptual changes in the relationship of an object to an underlying framework. A similar implication derives from the work of Werner, Wapner, and their colleagues, who offer similar explanations—hinging on postulated changes in equilibrial states—to account for both autokinetic motion and changes in egocentric spatial frameworks (Wapner and Werner, 1957; Werner and Wapner, 1949). These positions imply that autokinetic motion should be directly associated with experienced displacement of the autokinetically moving stimulus.

Phenomenological reports of paradoxical movement strongly challenge these views of the nature of the autokinetic phenomenon. This challenge has been made explicit in the work of Gregory and Zangwill. In an experiment that induced autokinetic motion as an aftereffect of selective fatigue of the eye musculature, Gregory and Zangwill obtained reports of 'paradoxical' motion. That is, motion appeared to take place without a corresponding shift of the apparent locus of the fixated light source. On the basis of these reports they conclude that "the autokinetic phenomenon . . . [is] caused by a temporary derangement of the neural systems directly concerned with the visual registration of movement" and is "in no wise due to faulty interpretation of the position of an object with regard to a 'framework,' perceptual or conceptual" (1963, p. 258).

It is unfortunate that so sweeping a conclusion is based only upon phenomenological reports of paradoxical movements. Although such reports are instructive, they are not conclusive. The issue seems to warrant a more rigorous experimental approach. Accordingly, the present experiments further investigated the question of whether motion and displacement are related. Rather than relying on phenomenological reports, however, the apparent displacement–autokinetic motion relationships were directly measured.

Experiment I investigated both apparent visual movement and apparent displacement of a luminous target under the conditions used by Gregory and Zangwill to produce autokinetic motion. Experiment II investigated changes in egocentric localization (apparent displacement) measured by nonvisual means under the same conditions.

EXPERIMENT I

Method

—Subjects—Twenty male subjects, all undergraduates at Yale University, served as observers.

—Method of inducing autokinetic motion—Two of the conditions for inducing autokinetic motion employed by Gregory and Zangwill (1963, p. 253) were used here. Subjects were instructed to hold their eyes 'hard over' to the right (or left) for a period of 30 sec. This is assumed to introduce an asymmetrical fatigue to the eye musculature. Upon return of the eyes to a normal straight-ahead gaze, measures of autokinetic motion were taken, presumably influenced by the aftereffect of this selective fatigue. In addition to the hard-over conditions, there was a third condition, in which the subjects held their gaze rigidly 'straight ahead' for the 30-sec adaptation period.

—Measurement of autokinetic motion and displacement—The type of measurement of autokinetic motion employed here closely resembles the 'tracking' method proposed by Bridges and Bitterman (1954). In this method the subject's task is to keep an object in a stable position by telling the experimenter how to move the object to its initial position after a movement has occurred. Thus subjects are required to 'compensate' for apparent movements by a real movement in the opposite direction.

This method has been criticized by Conklin (1955) as an impure test of autokinetic motion, since it is susceptible to the interactive effects of real movement (the compensatory movements of the object) on autokinetic movements. The intent of the present paper, however, was not so much to provide a complete and full description of autokinetic motion per se but, rather, to investigate the question of whether autokinetic motion is connected with spatial displacement. Since most the variance introduced by the tracking method would tend to be in a conservative direction and would not produce artifactual positive results, the tracking method was opted for here.

Spatial displacement was measured in a manner analogous to that used for the measurement of autokinetic motion. Subjects instructed the experimenter to adjust a luminous line to a position 'straight ahead, neither to the right nor to the left.'

—Procedure—All testing was done in a dark room. The subject was seated 8 ft from the stimulus array, with his head held rigidly in place by means of a headrest attached to his chair. With room lights on (but with the stimulus array shielded by means of a cardboard barrier), he was required to hold his eyes 'hard over' to the right (or left) or 'straight ahead' for a period of 30 sec. After the fatigue-induction period, the subject was instructed to close his eyes and to return them to (or keep them at) a normal straight-ahead position. When he signaled that his eyes were straight ahead, the room lights were turned off and the subject opened his eyes.

A dim vertical luminous bar, 1 in. wide by 6 in. long, which was placed objectively in the median sagittal plane, was observed.

The subjects were tested under two report conditions: tracking (the movement measure) and localization (the displacement measure). In the *tracking condition* the subject was told that the bar might move from its original position, and that any object motion in the right-left dimension was to be compensated for by a 'countermove' to bring the object back to its original position. In the *localization condition* he was told to adjust the bar until it was 'straight ahead' of him. Both tracking and localization measurements were initiated at eye opening and were continued for a period of 20 sec thereafter. If the subjects in the localization condition finished their adjustments prior to the 20-sec mark, they were encouraged to keep looking and refine their judgments.

A repeated-measures design was employed. Each subject was exposed to each of the three eye-fatigue conditions within each of the report conditions. Three minutes separated each condition of eye fatigue; 10 min separated each of the report conditions. The order of the report conditions was varied by a 2×2 latin square, and within each report condition the order of fatigue conditions was varied in a 3×3 latin-square design. The measures used in analysis were deviations in centimeters of the final position of the bar from its initial position.

Results

The results of the overall analysis of variance indicate that there were significant shifts dependent on direction of eye fatigue ($F = 36.23$, $df = 2/108$, $p < .01$). For eyes-right fatigue, there was a shift to the right (+18.8 cm), and for eyes-left fatigue there was a shift to the left (−9.9 cm). With eyes straight ahead, mean location was not significantly different from zero deviation (−0.29 cm). All differences between means were significant at the $p < .01$ level. There were no significant interactions involving the distinction between the tracking and localization measures.

The tracking measure is a measure of apparent motion in the right-left direction. For example, if a stimulus appears to move to the right, it is compensated for by a tracking movement to the left in order to maintain a constant position. The finding, then, that tracking was to the right (left) after the eyes-right (eyes-left) adaptation condition indicates that apparent movement was experienced as being toward the left (right). In this experiment, that is, it was found that autokinetic motion occurred in a direction opposite to previous eye turning. This finding, as far as it goes, agrees completely with that reported previously by Gregory and Zangwill (1963, pp. 253–254).

Although there is agreement between studies on the reported direction of autokinetic motion, there is no agreement on whether this motion is associated with displacement. The present study found no differences between measures of tracking and of localization. If motion were 'paradoxical,' and thereby independent of experienced displacement, we would expect either that both measures would fail to show shifts dependent on fatigue condition, or that the tracking method would show shifts (since it could be construed as a measure of motion whether real-appearing or paradoxical) while the localization measure would not. The fact that both measures showed entirely comparable shifts dependent on eye-adaptation preconditions suggests that the motion obtained here was associated with a corresponding displacement.

The relationship between motion and displacement was tested by other means. The magnitude of directional changes was assessed for each subject by computing an 'eyes-right minus eyes-left' difference score for both the tracking and the localization condition. These measures were then correlated. A significant ($p < .01$) positive correlation ($r = .54$) indicates that motion and localization were, under the conditions of the present experiment, related.

EXPERIMENT II

The findings of relationship obtained in Experiment I are subject to several limitations as long as measurements of motion and localization are obtained within the same (visual) modality. The first limitation stems from a problem of measurement; the second from a problem of interpretation.

With the measures employed in Experiment I there is the possibility that since the subjects were provided no response measure which would allow them to distinguish between 'real-appearing' and 'paradoxical' autokinetic motion, they used the response measures which were provided in a way that did not distinguish these two possibilities. While this argument is reasonably applicable to the tracking method; it seems less applicable to the localization measure, where there is a less ambiguous indication of the task.

The interpretational limitation of Experiment I hinges on the type of dependency relationship between motion and localization. Is it that an object appears to move—and hence changes its location with respect to a fixed spatial coordinate? Or is it that the coordinate is induced to change—and hence the object appears to move?

Both of these limitations suggested an experiment where measurement of location shifts is nonvisual and therefore decoupled from apparent object motion. Experiment II met this requirement by measuring shifts in egocentric localization by means of the nonvisual method of 'pointing straight ahead' in the absence of an object, but under the conditions that produced autokinetic motion in Experiment I.

Method

—Procedure—Experiment II was a replication of the localization condition of Experiment I, with two major exceptions. The first was that after eye fatigue (right; left; straight-ahead) was induced and the eyes returned to a normal straight-ahead position, the subjects kept their eyes closed. The second exception was that a nonvisual determination of the straight-ahead was substituted for the visual means employed in Experiment I. Here, the subject, his eyes closed, was required to extend his arm toward a measuring surface (a ruler located in the frontal-parallel plane) and to point 'straight ahead.' After a single eye-fatigue period, three determinations of the straight-ahead were made in quick succession.

—Subjects—The subjects were 24 right-handed Yale undergraduate men.

Results

There was a significant effect of eye-adaptation condition on nonvisual determination of the straight-ahead ($F = 5.56$, $df = 2/24$, $p < .05$). The effects obtained roughly parallel those obtained when visual determinations were used. Location of the straight-ahead after the eyes-right condition was significantly ($p < .05$) to the right (+5.20 cm) of the location of the straight-ahead after either the eyes-left (+4.41 cm) or the eyes straight-ahead (+4.06 cm) conditions. However, there was no significant difference between the 'eyes left' position of the straight-ahead and the 'eyes straight ahead' position of the straight-ahead. All measures are displaced to the right of the objective straight-ahead.[1]

The general directional parallels between visual and nonvisual determinations of the straight-ahead are of singular interest. They provide evidence of a general organismic change that is not limited to 'error signals from the visual monitoring system' resulting from conditions of eye adaptation.

Although there is evidence for a generalized effect, an examination of intersubject consistency suggests that the effect of eye adaptation was stronger for the visual measures.[2] In Experiment I, 90% of the subjects showed an effect consistent with the mean; here in Experiment II only 70% of the subjects showed this effect.

DISCUSSION

Experiments I and II both provide evidence in support of the notion that, under the conditions used by Gregory and Zangwill, measures of motion and measures of displacement are indeed related.

Moreover, the data of Experiment II suggest that the relationship holds beyond the purely visual sphere. Both the fact of intersensory interaction (a fact attested to in Experiment I, where fatigue of eye musculature led to shifts in visual localization and movement, and in

Experiment II, where tactual kinesthetic localization was similarly affected) and the suggestion of generalized changes in frameworks for localization—both of these findings fit well within the interpretational framework of sensoritonic field theory (Werner and Wapner, 1949; Wapner and Werner, 1949). Rather than attempting a detailed explanation of these findings within the context of sensoritonic field theory, let it now suffice to point out that the dual notions of organismic interrelatedness of systems and generalized equilibrial changes within an organismic field suggest the type of results obtained here. This interpretational framework clearly contrasts with the one adopted by Gregory and Zangwill, who attempt to locate the autokinetic effect within a segregated subsystem of the organism, namely, the eye-movement monitoring system.

Although we have provided clear evidence of the relationship between changes in framework and direction of perceived motion, the argument of this paper should not be construed as a claim that autokinetic motion is determined only by changes in object-framework relationship. There is considerable evidence, some based on phenomenological report (Moustgaard, 1963) and some based on direct measurement (Glick, Wapner, and Werner, 1965; Glick and Wapner, 1967), for the occurrence of 'paradoxical' relationships between motion and location shifts. It is as yet unclear whether all of those phenomena described under the umbrella term 'autokinetic motion' are related in terms of underlying process. It is quite possible that a physicalistic definition of autokinetic motion as 'apparent motion in the absence of real motion' serves only to obscure critical distinctions and groupings that may be made on the basis of process. What is needed are systematic studies of motion-displacement relationships under a wide variety of 'dimensionalized' situations. Should differential relationships depend on situational dimensions, then some first steps could be taken toward developing *theories* of the *motion processes*.

Notes

This research was supported by Grant HD-03947 from the National Institute of Child Health and Human Development to the author, who was then at the Institute of Child Development at the University of Minnesota. His current address is at the City University of New York, Graduate Center, 33 West 42nd Street, New York, New York 10036. Received for publication June 27, 1970.

1. This is most likely due to the fact that the right hand was used to point. Informal observations indicated that left-hand pointing led to a systematic displacement of all measures to the left.

2. Comparison of magnitudes of deviation for visual and nonvisual measures is a tenuous business because of the proximal nature of the pointing measure and the distal nature of the visual measure. Computation of visual angle or of a comparative 'pointing angle' is difficult because of the difficulty of fixing an apex for the pointing measure. Accordingly, this mode of comparison was not used here.

References

Bridges, C. C., and Bitterman, M. E. 1954. The measurement of autokinetic movement. *American Journal of Psychology* 67:525–529.

Brosgole, L. 1967. Note on perceived movement and apparent displacement. *Acta Psychologica* 27:233–235.

Carr, H. A. 1910. The autokinetic sensation. *Psychological Review* 17:42–75.

Conklin, J. E. 1955. A note on the measurement of autokinetic movement. *American Journal of Psychology* 68:144–145.

Glick, J. A., and Wapner, S. 1967. The relationship between perceived movement and apparent displacement: A reply to Brosgole. *Acta Psychologica* 26:236–240.

Glick, J. A., Wapner, S., and Werner, H. 1965. Some relations between autokinetic motion and space localization. *Acta Psychologica* 24:41–48.

Gregory, R. L., and Zangwill, O. I. 1963. The origin of the autokinetic effect. *Quarterly Journal of Experimental Psychology* 15(4):252–258.

Moustgaard, I. K. 1963. A phenomenological approach to autokinesis. *Scandinavian Journal of Psychology* 4:17–22.

Wapner, S., and Werner, H. 1957. *Perceptual development: An investigation within the framework of sensory-tonic field theory*. Worcester, Mass.: Clark University Press.

Werner, H., and Wapner, S. 1949. Sensory-tonic field theory of perception. *Journal of Personality* 18:88–107.

FORM AND ITS ORIENTATION: REEXAMINATION OF A CHILD'S-EYE VIEW

Lauren Harris and M. Joseph Schaller
Michigan State University

One hundred twenty kindergarten, second-, and fourth-grade children viewed bilaterally symmetrical geometric forms through a circular window on a turntable and were asked to make them right side up or upside down. The greatest number of nonvertical orientations were made for forms that resembled horizontally oriented English letters. Upside-down judgments were not the empirical converse of right-side-up judgments. The results are compared with those from an earlier study by Braine.

Lila Ghent Braine (Ghent, 1961) has described two studies of children's judgments of the orientations of certain geometric forms (which she called "nonrealistic" forms). The forms employed in her first study, and the only forms with which we were concerned in this work, are reproduced in the left column of Table 1, with the exception of an equilateral triangle. Her subjects ranged in age from four to eight years. The agreement among subjects was striking: for the four-year-olds, agreement of judgment of 'upside down' was greater than chance for 11 of the 16 forms. Indeed, for 6 of these forms the percent agreement was 95%. The older children's judgments agreed strongly for fewer forms and reversed in judgment for 2—forms 1 and 11.

Braine saw the most important problem raised by these findings to be that of "defining the characteristics" of the cards selected as upside down. Using responses of her younger subjects as the basis for her analysis, she suggested that a form was judged upside down when the "focal portion was in the lower half of the card." Braine tentatively assumed that figures eliciting significant concordance of judgment had one portion that "caught the 'attention,' or drew the eye, more readily than did other portions. The angle of [form 1], the point of intersection of [form 2] the rounded portion of [form 10], and so on, might be considered the focal portion. . . . When these portions of the figures were at the bottom, the cards were called upside down by the children" (1961, p. 181).

Recognizing that her designation of 'focal' had been post hoc, Brain tested different four- and five-year-olds with new forms for which she designated the focal points beforehand. These subjects' judgments supported her focal-point hypothesis.

The current study grew partly out of consideration of the method used in the Braine study, but primarily out of our previous interest in the relation between the development of spatial orientation and letter discrimination. We therefore used the forms from Braine's first experiment, since a number of them are either real letters or are letterlike.

We were concerned with two features of Braine's method. First, Braine used what we may call a 'binary-vertical forced-choice procedure' to obtain her judgments of orientation. Her subjects saw two cards at a time, both depicting the same form, one rotated 180 deg with respect to the other. In both orientations, moreover, the form was fixed vertically, meaning that it was symmetrical along its vertical axis but not along its horizontal axis (see Table 1, column 1). The subjects therefore were *not* shown the forms oriented either 'horizontally' (symmetrical along their horizontal axes but not along their vertical axes) or 'obliquely' (symmetrical along neither their vertical nor horizontal axes). Consequently the subjects had no chance to judge either of these orientations as the upside-down orientation.[1] And even if they had been allowed to choose from additional orientations, the placement of the forms within square frames might have produced a 'frame' effect, thereby influencing orientation judgments.

In other words, there are at least two kinds of cues that potentially can affect the perceived orientation or identification of a figure: cues *within* the figure itself (what Howard and Templeton, 1966, have called "intrinsic factors" such as axes of symmetry, taper, texture and light gradients, and Braine's 'focal point'); and cues from *without* the figure ('frame cues'). Braine's procedure may have confounded these two factors, and it is possible that the particular responses made by her subjects were also to some degree affected by frame cues.

We therefore presented the forms in an apparatus specially designed to eliminate all constraints on range of judgments and to reduce frame-of-reference cues. The question we were asking, then, was whether the substitution of our procedure for Braine's binary-vertical procedure had a significant effect on the children's orientation judgments. Would the children, now with no constraint on their range of judgments and with the possibility of frame cues substantially reduced, choose to orient the forms a significant number of times in orientations other than Braine's two vertical orientations?

Our second concern was with Braine's instructions. Her subjects did not make direct choices of the upright forms: they were asked to pick the form that was upside down or 'wrong' rather than the form that was upright. Then, from their choice of the upside-down, wrong form, Braine inferred that they would judge the other form (the 180-deg rotation) to be right side up. But are 'upside down' instructions in fact the empirical converse of 'right side up' instructions? To answer this question, we compared the two kinds of instructions directly.

METHOD

—Apparatus—The apparatus consisted of a 60 by 79 cm Masonite board covered with enough coats of white paint to eliminate all perceivable grain. In the center of this board a circle (9-cm diameter) was cut, through which the forms could be viewed. The forms were inserted into a Plexiglass holder, which was mounted on a turntable. The turntable was bolted vertically to a wooden door that was attached on one side to the back of the Masonite board. The wooden door, when closed against the Masonite board, brought the center section of the Plexiglass holder nearly flush against the Masonite board so that the form appeared through the circular window. The form that the subject saw thus appeared in the middle of the circular window. The effect was to eliminate completely any frame-of-reference cues that depended on the orientation of the form and could therefore affect the subject's judgment of a form's upright orientation. The apparatus was clamped to a table so that the front of the Masonite board was flush with the front of the table. The board actually was tilted back approximately 5 deg, not enough to induce any significant change in perceived slant of any of the forms. (In any event, the subjects tended to hunch forward while rotating the forms, thus probably compensating for whatever mild distortions of slant might have been induced.) The apparatus was placed in a small spare room in the school.

We had found, in a pilot study, that a number of children peeked through the hole to watch the experimenter insert the form in the Plexiglass holder. Since we expected that seeing this placement might influence the subject's subsequent judgment of the form's orientation, we masked the hole while inserting the form. We also gave the turntable a hard spin before exposing the form so that the form was rotating when the subject first saw it. The orientation of the form when it came to rest thus varied randomly from form to form and from subject to subject.

The subject rotated the form directly by moving any part of the Plexiglass container that showed through the circular window (roughly as one might rotate a phonograph record with one's fingers, except that the turntable was presented in nearly vertical orientation). The turntable was so adjusted as to permit easy rotation by all subjects. The apparatus thus eliminated all constraints on the subject's range of judgments.

All orientations were recorded to the nearest degree by means of a pointer attached to the back of the turntable, which rotated about a 360-deg scale. Orientations later were categorized into 16 sectors, each 22 deg 30 min wide (11 deg 15 min on either side of the cardinal point). Pilot testing indicated that even our youngest subjects could make adjustments within an arc of 22 deg 30 min. Thus if our procedure were to yield a significant number of oblique judgments, we would be able to say that they were not artifacts of the children's inability to make fine-enough motor adjustments of the turntable. The children did not have to align the form perfectly upright to qualify for a vertical score.

—Subjects, procedure, and instructions—There were 120 children tested: 20 boys and 20 girls at each of three levels: kindergarten, second and fourth grades. Half the children in each level were given 'right side up' instructions, the other half were given 'upside down' instructions. We ascertained the children's understanding of our instructions through pretesting with a set of ordinary toys (cars, trucks, dollhouse tables and chairs). In the upright instruction condition of this pretesting, the experimenter said: "I'd like to know how children can know how to put things so that they're right. Look at these things here [pointing to the toys, all of which were in upside-down and sideways orientations on a table just to the subject's right]. Can you turn them to make them right, the way they're supposed to be?" The experimenter waited until the subject had turned all the toys to their proper, right-side-up, orientation.

All the children performed perfectly and quickly. This fact gave assurance that, at least for ordinary toys, the instructions to 'make them right, the way they're supposed to be' were equivalent to 'make them upright.' We should note that we were reluctant to instruct the children merely to make the objects upright or right side up, since our pilot work suggested that some kindergartners were not familiar with these words.

The pretesting in the upside-down instruction condition was identical except that some of the toys now lay right side up, others still lay on their sides, and the child was asked to turn the toys "to make them upside down, to make them wrong." Again all the children performed perfectly and quickly.

Immediately after pretesting in both instruction conditions, the children were shown how to rotate the forms in the apparatus and were again given the pretest instructions, this time for the forms, which the experimenter called "pictures." Thus the instructions for children in the upright (or upside-down) condition continued: "Now, what about making *pictures* of things right, the way they're supposed to be [or "wrong, what about making pictures upside down"]? I'll put a picture in this little window [indicating the circular window]. See, the window moves around, you can move it with your fingers. You can turn it this way, or this way. [The experimenter demonstrated, and then let the subject move the window with his fingers in each direction.] Now, I'll put a picture in for you to look at [inserting the first form]. You turn the picture until it's right—the way it's supposed to be, just the way you made these things right [or "until it's wrong, until it's upside down, just the way you made these things upside down"]." The experimenter pointed to the toys as he spoke these instructions. The subject then was permitted to turn the form until he stopped—presumably, then, until he found an orientation that he judged to be upright (in the upright instruction condition) or upside down (in the upside-down instruction condition). The experimenter then recorded the pointer's reading for that form, then inserted the second form, reinstructed the subject, and so on through the entire series.

All subjects were shown all of the forms depicted in column 1 of Table 1. The equilateral triangle Braine used was excluded because it has three equivalent axes of symmetry rather than one. This fact would have posed some problems for the computer analysis of the data. All forms were 5–6 cm in size. The 15 forms were presented in five different orders, with an equal number of subjects in each age × sex group assigned randomly to each order. The five orders were constructed so that every group of three successive forms (1 2 3, 4 5 6, and so on) occurred once and only once in any position within the order.

The total testing time was approximately 10 min for each subject. Most of the subjects made each of their judgments within 10–15 sec. Those few who took longer on a few trials seemed to be playing with the rotating window. When this happened, the experimenter simply reminded the child of his task. Each child was told that he had "done a good job" after the session was completed.

RESULTS

The first question, about the effects of providing only the two vertical choices, may be answered by consideration of those 60 children who received the 'upside down' instruction (as had Braine's subjects); in other words, those subjects who differed from Braine's only with respect to the apparatus used.

If a binary-vertical procedure like Braine's does not significantly constrain the subjects' range of judgments, we would expect that our subjects, though free to make orientation judgments in nonvertical categories, would restrict their choices to the two vertical categories.

The first analyses, then, consisted of seeing how *well* the total distribution of scores across forms for each grade, and the distributions for each form individually for each grade, *fit* the hypothesized binary-vertical distribution. The Kolmogorov-Smirnov goodness-of-fit test was used for this analysis.

The distributions and results of tests on total scores are depicted in Figure 1. Categories 1 and 2 represent the two 'vertical' categories used in Braine's binary procedure. Categories 3 and 4 represent 'horizontal' orientations, and category 5 represents any oblique orientation. For the total distributions for each grade, the percentage of orientations into *any* nonvertical cagetory (i.e., summing across categories 3, 4, and 5) was 27% for kindergartners, 11% for second graders, and 23% for fourth graders. These percentages were sufficiently large to indicate that a binary-vertical presentation would have significantly constrained the subjects' choices ($p < .01$) for each grade.

What of the judgments of the 60 children who received the 'upright' instructions? This condition, we note again, differed from Braine's in both apparatus and instructions. Figure 2 depicts the distribution of judgments in the upright instruction condition. The percent orientation into some nonvertical category (categories 3,

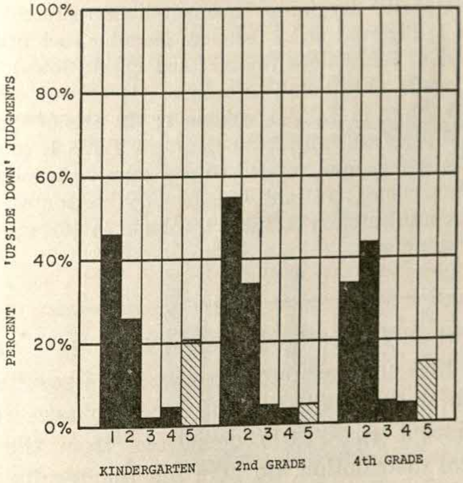


Fig. 1. Distribution of judgments of 'upside down' for all forms combined for each of the three grades (20 subjects per grade, each subject making 15 judgments; total $N = 300$, each grade). Orientation categories: (1) axis of symmetry vertical, form rotated 180 deg from orientation shown in Table 1, column 1; (2) axis of symmetry vertical, form oriented as shown in Table 1, column 1; (3) axis of symmetry horizontal, form rotated 90 deg clockwise from orientation in Table 1, column 1; (4) axis of symmetry horizontal, form rotated 90 deg counterclockwise from orientation in Table 1, column 1; (5) axis of symmetry in any oblique orientation.

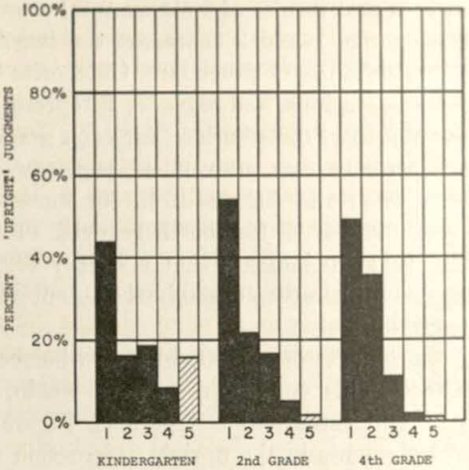


Fig. 2. Distribution of judgments of 'upright' for all forms combined for each of the three grades (20 subjects per grade, each subject making 15 judgments; total $N = 300$, each grade.) Categories 1–4 in this figure represent orientations rotated 180 deg from the orientations represented by the categories in Figure 1; Figures 1 and 2 therefore resemble each other to the extent that the two sets of instructions (upright and upside-down) produce empirically converse results. Orientation categories: (1) axis of symmetry vertical, form oriented as shown in Table 1, column 1; (2) axis of symmetry vertical, form rotated 180 deg from orientation shown in Table 1, column 1; (3) axis of symmetry horizontal, form rotated 90 deg counterclockwise from orientation in Table 1, column 1; (4) axis of symmetry horizontal, form rotated 90 deg clockwise from orientation in Table 1, column 1; (5) axis of symmetry in any oblique orientation.

4, and 5) was 41% for the kindergartners, 24% for the second graders, and 15% for the fourth graders. The results of goodness-of-fit tests on the total distribution for each age-group again indicated significant ($p < .01$) deviations from the hypothesized binary-vertical distribution. So here too the results indicate that, had only the two vertical response categories been provided, the children's judgments would have been constrained significantly.

The conclusion can be drawn that insofar as total scores for the three age-groups are concerned, a binary procedure would significantly constrain the range of judgments whether the instructions are to place the forms upright or upside down. But are the instructions equivalent? In this case, this question amounts to asking whether the *character* of the constraint for the upright instruction condition is the same as that for the upside-down instruction condition. Figure 3 condenses Figures 1 and 2 into a side-by-side comparison of the distribution of judgments, combining age groups for each instruction condition. Results of a goodness-of-fit test indicated that the distributions differed significantly ($p < .01$) from one another. The difference stems primarily from the distribution of judgments in the nonvertical categories (categories 3, 4, and 5). The vertical categories (1 and 2 combined) were not significantly different between conditions. The total percentages of nonvertical judgments (horizontal and oblique combined) are roughly equal, but the upside-down instructions induced about twice as many oblique judgments as horizontal judgments, while the upright instructions induced about three times as many category 3 judgments as either category 4 horizontal judgments or oblique judgments. The instructions therefore are *not* equivalent, at least in terms of the total distribution of judgments. Asking a child to make these forms 'upside down' or 'wrong' is not the empirical converse of asking him to make the same forms 'right, the way they're supposed to be.'

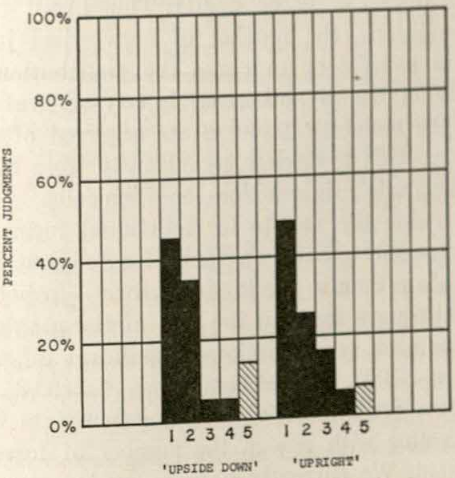


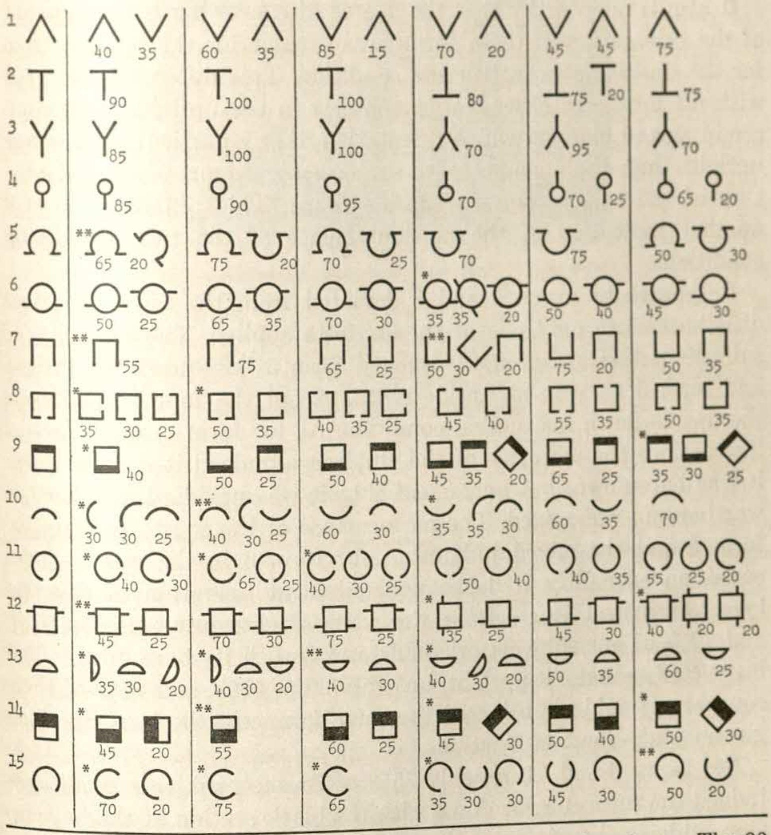
Fig. 3. Distribution of judgments for all forms for all grades combined (Figures 1 and 2 condensed; 60 subjects making 15 judgments each = 900 judgments total for each distribution). Judgments represented by categories 1–5 are the same for 'upside down' and for 'upright' as in Figures 1 and 2 respectively; these two distributions therefore resemble each other to the extent that the two sets of instructions (upright and upside-down) produce empirically converse results.

But the picture becomes considerably more interesting when distributions for the individual forms are studied. Distributions for 10 of the 15 forms included a sufficiently *small* number of horizontal and oblique responses to indicate that the provision of only the two vertical categories would not have been a significant constraint. For these forms, the great majority of 'upright' judgments placed the forms in 'vertical' orientation (symmetrical sides on left and right). But for the 5 remaining forms (forms 10, 15, 13, 8, 11), the provision of the binary choice would have lent a distorted picture of the kind of upright judgments the children made. The significantly frequent use of the nonvertical categories occurred in only 30 of the 90 form × grade × instruction-condition cells. The 90 cells are depicted in Table 1. The 30 deviant cells are marked with asterisks.[2] In each cell of the table, only those orientations of a form that accounted for at least 20% of the judgments made are shown. The percentage of the choice of these orientations appears as a subscript to each form. The forms are numbered (1–15) in order of increasing number of significant deviations from the hypothesized binary-vertical distribution (as determined by the goodness-of-fit test, $p < .05$ or $< .01$). Thus in all grades in both instruction conditions, the number of nonvertical judgments for forms 1–4 was insufficient to cause the distributions to deviate significantly from the hypothesized binary-vertical distribution. For form 15 the nonvertical judgments occurred often enough to cause the distribution to deviate significantly from binary-vertical in five of the six grade × instruction-condition cells.

Before discussing the results for individual forms, we note in Table 1 the generally greater strength of agreement in the second and fourth graders than in the kindergartners—in both instruction conditions. Chi-square tests on the sum of the greatest frequency in each cell for each grade disclosed significant differences ($\chi^2 = 28.9$, $df = 2$, $p < .001$, upright condition; $\chi^2 = 8.63$, $df = 2$, $p < .02$, upside-down condition). This finding contrasts with Braine's report of a decline with age in the number of forms that elicit strong agreement. We currently have no explanation for this apparent discrepancy.

The scores for forms 1–4, for which there were no significant deviations from the binary-vertical, are in substantial agreement with Braine's findings. The form 1's adjudged upside down shifted regularly from apex-down to apex-up across age. The converse occurred in the upright condition, with a shift from apex-up to apex-

Table 1. Popular (20% or more of judgments made) orientations for each form



	'Upright' instructions			'Upside down' instructions		
Form	Kinder-gartners	Second graders	Fourth graders	Kinder-gartners	Second graders	Fourth graders
1	40 35	60 35	85 15	70 20	45 45	75
2	90	100	100	80	75 20	75
3	85	100	100	70	95	70
4	85	90	95	70	70 25	65 20
5	** 65 20	75 20	70 25	70	75	50 30
6	50 25	65 35	70 30	* 35 35 20	50 40	45 30
7	** 55	75	65 25	** 50 30 20	75 20	50 35
8	* 35 30 25	* 50 35	40 35 25	45 40	55 35	50 35
9	* 40	50 25	50 40	45 35 20	65 25	* 35 30 25
10	* 30 30 25	** 40 30 25	60 30	* 35 35 30	60 35	70
11	* 40 30	* 65 25	* 40 30 25	50 40	40 35	55 25 20
12	** 45 25	70 30	75 25	* 35 25	45 30	* 40 20 20
13	* 35 30 20	* 40 30 20	** 40 30	* 40 30 20	50 35	60 25
14	* 45 20	** 55	60 25	* 45 30	50 40	* 50 30
15	* 70 20	* 75	* 65	* 35 30 30	45 30	** 50 20

Note: Subscripts indicate percent judgments in orientation shown. The 30 cells marked with asterisks are those cells in which distributions of orientation judgment differed significantly (*$p < .01$; **$p < .05$) from an hypothesized binary-vertical distribution. Oblique forms indicate *any* oblique orientation. $N = 20$ per cell.

down across age. Apparently, for the younger children intrinsic factors such as taper and stability affected their judgments, such factors being overpowered in influence for the older children, who regarded the form as the letter V. (It also is possible that the younger children thought the form an A without the crossbar.)

For forms 2, 3, and 4, 85%–100% of the children oriented them, in the upright condition, as T, a Y, and what might be called a 'balloon on a string.' In the upside-down condition, 65%–95% of the children turned the same forms 180 deg. These findings also agree substantially with Braine's.

It also is noteworthy that the degree of concordance of judgment of the orientation of these forms was greater for the upright than for the upside-down instruction condition. This difference occurred with *all* forms. In other words, subjects in the upright instruction group agreed more on which orientation of any particular form was upright than their counterpart subjects agreed on which orientation of the same form was upside down. This difference is still another indication of the nonequivalence of the two instruction conditions.

Responses to forms 5 and 6 deviated from the binary-vertical distributions in only one of the six tests applied. The majority of subjects across age-groups adjudged form 5 bars-down as upright and form 5 bars-up as upside down, except the fourth graders in the upside-down instruction condition. As for form 6, with upright instructions the majority of the children adjudged it upright when it was bars-down, but agreement that it was upside down when it was bars-up was mixed. The performance of the fourth graders for form 5 and the mixed judgment for form 6 in the upside-down condition exemplify well the nonequivalent operation of the two types of instructions *within* the vertical categories: 50%, 65%, and 70% of our subjects oriented form 6 with its bars down when instructed to make it upright, but only 20%, 40%, and 45% of their counterpart subjects oriented the same form with its bars up when given upside-down instructions.

For forms 9 and 14, roughly 75% of Braine's subjects at all ages judged the forms upside down with the black section at the bottom. Our subjects showed no such strong agreement, as is evident from Table 1. Indeed, in some cases, their strongest judgments of both 'upright' and 'upside down' were of the forms in the same orientation.

We now consider forms 7 and 8, 10 and 11, and 15, all of which, when oriented horizontally, resemble—to varying degrees—the letter C. In Figure 4 these five forms are arranged in the order in which the subjects in the upright condition oriented them, as they presumably would have oriented a real letter C. Though the shapes

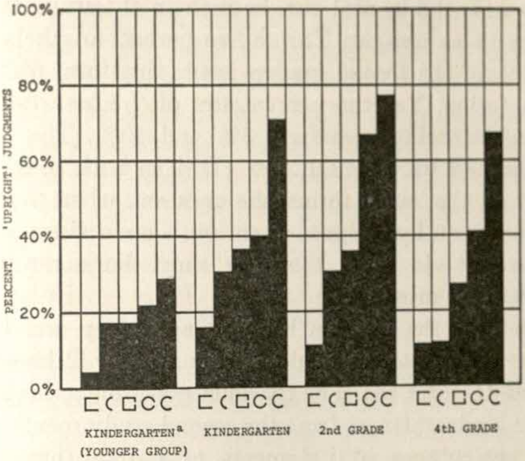


Fig. 4. Distribution of C-like (horizontal, open to right) judgments, in the upright instruction condition, of five forms resembling the letter C ($N = 24$ for Kindergarten[a]; all other Ns $= 20$). Forms are arranged in increasing order of resemblance to letter C, by post hoc analysis.

of the distribution were different for each group, the ordering of the five forms was the same. (We were hoping for some similarity but hardly expected *identity* of ordering across three age-groups.) Figure 4 also presents the same judgments for a slightly younger group of kindergartners tested in a second study (Kindergarten[a]). Again the ordering is the same. We repeat that this ordering is post hoc. We currently are trying to get independent assessments of children's judgments of the similarity of these forms to the letter C—and also to the letter U—to see whether their judgments jibe with the orderings shown here.

It is surprising how rarely the children who received *upright* instructions (see Figure 4) placed these forms in what might be called 'backward-C' orientations, that is, open-to-left orientations—surprising in view of the frequency with which young children reverse letters in reading and writing. For the C itself (form 15), the percentages of backward-C orientations were zero for all three grades (and also for the young kindergartners); for the four remaining forms, the percentages were zero for the second and fourth graders, and averaged only 10% (range 5%–15%) for the kindergartners.

This same rejection of a 'backward' orientation for a form that is letterlike when horizontal can be seen in the ways the children judged form 13 as upright. Thirty-five percent of the kindergartners and 40% of the second graders made this form upright in the D-like orientation. Yet the percentages of 'backward-D' orientations were substantially smaller: 5% and 10%. The percentages for fourth graders were equally low (15% in both instruction conditions. For all the 'open' forms, the children, when they did align the forms horizontally, aligned them with open sides facing right rather than left—in fact, the way most horizontally oriented English letters are oriented.

The pattern of the relation between age-group and the percent of letterlike orientations is nonlinear for all four C-like forms (except the true C, form 15). For the four age-groups, beginning with the younger kindergartners from the second study mentioned above, the mean percentages of judgments of C-like forms in C-like orientations were 15%, 31%, 35%, and then down to 21% for fourth graders.

Perhaps the shape of these distributions reflects the children's increasing familiarity with the critical features of letters, a familiarity which, through the second grade, increases the likelihood of their reponding to the forms as though they indeed were letters. By fourth grade, however, perhaps the children have learned the features sufficiently well—that is, they can discriminate the letters so skillfully—that the letterlike forms now are not letterlike enough. (We reached the same conclusion on the basis of a previous study of children's same-different judgments of letterlike forms; Schaller and Harris, 1969.)

The pattern of responses to the C-like forms also illustrates our proposition that the two instruction conditions are nonequivalent for certain forms. Only 1%, 6%, and 4% of the kindergartners, second graders, and fourth graders respectively turned the C-like forms 180 deg from the C-like orientations (turned them so that their open sides were facing left) when asked to make them upside down. Were these forms like letters that stand upright when oriented vertically (such as the T, V, or Y), the children should have had no trouble turning them upside down. But when a form resembles a letter that is normally oriented horizontally, the problem of specifying what is upside down becomes much more difficult. A letter C turned upside down by rotating it 180 deg is in fact a backward C. The data indicate that the children rejected backward C's as candidates for upside-down C's. And, of course, a C turned upside down by up-down reversal is identical to the upright C. These facts about rotational transformations of letters one might not expect young children to apprehend. So, instructed to make the letter or letterlike form upside down, the children did all they could—they simply aligned it in some orientation other than backward- or forward-horizontal; thus in the present case, in either a vertical or an oblique orientation (as shown in Table 1).

DISCUSSION

In summary, these data suggest the need for caution in referring to these forms as 'nonrealistic.' If they do not resemble real objects, many of them do resemble—and apparently are responded to as—English letters. Therefore, any attempt to understand the bases on which they are judged as having orientations in space cannot be limited to an analysis of intrinsic factors. Under some conditions, certainly, an analysis in terms of intrinsic factors such as focal point, distribution of mass, lines of symmetry, taper, texture and light gradients, may account fully or primarily for children's (or adults') judgments of the orientation of such forms.

But if the form bears any resemblance to an English letter, then the intrinsic factors may be balanced or even overpowered in influence by the fact that the form is recognized as a letter. Thus while symmetry, 'balance' or 'stability,' and distribution of mass predispose form 15 to be oriented facing down, its identification as the letter C will cause it to be oriented facing right. This notion implies difficulty for any developmental analysis of the role of intrinsic factors in the discrimination of orientation, because age undoubtedly is correlated highly with increasing familiarity with the alphabet. The analysis is further complicated if we are correct in postulating a *nonlinear* relation between age and the frequency of letter-appropriate orientation judgments of forms that are merely *letterlike*.

Another factor that seems to influence judgments of orientation is symmetry. All the forms used, in both Braine's experiment and ours, have one axis of symmetry and usually were oriented with the axis vertical. The exceptions were the forms that resembled letters whose orientations are horizontal. Knowing that the axis is oriented vertically does not, however, tell us which end of the axis is the top and which the bottom. For judgment of top and bottom, intrinsic factors other than line of symmetry, factors such as distribution of weight, taper, and Braine's 'focal point,' must be considered.

The data also suggest that an analysis of the factors which govern a form's designation as upright may or may not help us to understand the bases for the designation of the upside-down orientation of the same form. Whether the form is letterlike and whether the letter is normally oriented vertically or horizontally seem to be very important factors.

There are other factors about which we can only speculate at this point. For instance, some objects or representational forms have one unequivocal upright orientation (house, man, car), while others can be oriented upright in more than one position (key, pencil, book, fork, eraser). It would appear that this characteristic, whether the object is 'monoriented' or 'polyoriented,' would affect the degree to which one could successfully infer the object's right-side-up orientation from its adjudged upside-down orientation, and vice versa. It probably is acceptable to infer that the monoriented house or car would be judged right side up when turned 180 deg from the position in which it is judged upside down, but this inference is suspect for the polyoriented pencil or fork. A similar objection can be brought, a fortiori, to the assumption that the choice of 'upside down' for a certain orientation of a *non*representational form implies that its opposite is right side up. Only then, for forms 2, 3, and 4, the only unequivocally monoriented forms used, would we expect an accurate inference of one orientation from the other.

Whether or not a realistic form is monoriented might not depend totally on the extent of agreement about its 'upright' orientation. For instance, assume that 100% of a sample of adults agreed that an ordinary sharpened pencil was upright with its writing point up, its eraser end down. (In fact, we have found that the agreement is not 100%, but it is useful to assume perfect agreement for purposes of this illustration.) Assume also that the same persons agreed that a man, a table, and a horse were upright when their heads were at the top, their feet at the bottom. (In fact, all our adults did agree to this.) On this basis, the pencil, man, horse, and table could be said to be monoriented. But now rotate each figure 180 deg from its judged upright position and ask the adults whether each is now upside down. They agree with no hesitation that the man, horse, and table are upside down, but some of them feel "funny" describing the pencil as 'upside down' with its writing point at the bottom, its eraser at the top, for that is how they hold the pencil when they write. In other words, a pencil *can* be upside down, but with respect to some particular function rather than with respect to gravity.

This example, if accurate, portends even greater difficulty for a developmental analysis of the role of intrinsic factors in the discrimination of orientation. For if the forms are seen as representations of real objects and not representations of letters, we must now be concerned with the *character* of these objects. How are they normally oriented? Are their upside-down orientations 180-deg transformations of their right-side-up orientations? Do they have clearly defined orientations at all? How does the judgment of orientation of real objects themselves change with the age and experience of the observer? These are some of the questions we are trying to pursue in further work.

Notes

Portions of this research were reported to the Psychonomic Society, November 7, 1969, Saint Louis, Missouri. The authors are grateful to the principal, staff, and students of the Spartan Public School, East Lansing, Michigan, and the Midway Public School, Holt, Michigan, for their generous cooperation in this study. Thanks also are due Daniel C. Gutkin and Gary C. Johnson for their assistance, and Nancy Hammond for her criticisms. The research was supported in part by a National Science Foundation biomedical grant and a Michigan State University faculty grant to the first author. Dr. Schaller is now at the University of Wisconsin, Madison. Received for publication July 2, 1970.

1. Braine, however, mentioned obtaining substantially similar results with subjects given "the pile of cards, in random order and orientation," with instructions "to place each card in the right-side-up position" (1961, p. 184). The five- and six-year-olds however, "did turn some of them so that the figures were 'symmetrical about the horizontal axes,' . . . primarily . . . the circular figures with an open portion, and the closed crescent." In these instances the children spontaneously said, " 'looks like a D,' 'looks like a C' " (personal communication to LH, November 18, 1969).

2. Strictly speaking, comparing our data with Braine's is not possible for the 30 'deviant' cells. In other words, since the subjects' 'upright' judgments fell in vertical categories 1 and 2 for form 12 only 45% and 25% of the time, and resulted in a significant nonvertical distribution, we could not assume that, had only the two vertical categories been provided, there still would have been about twice as many category 1 judgments as category 2 judgments.

References

Ghent [Braine], L. 1961. Form and its orientation: A child's-eye view. *American Journal of Psychology* 74:177–190.

Howard, I. P., and Templeton, W. B. 1966. *Human spatial orientation*. New York: Wiley.

Schaller, M. J., and Harris, L. 1969. Children's same-different judgments of various transformations of letter-like forms using the method of paired comparisons: Effects of transformation and age of child. Paper presented to the Society for Research in Child Development, Santa Monica, California, March, 1969.

PATTERN COMPLEXITY AND RESPONSE TO DISTINCTIVE FEATURES

Robert Pasnak
Catholic University of America

Experiment I showed that subjects respond to distinctive parts of complex random forms but to the whole contour of simple random forms. Experiment II showed that the probability that a subject would respond to a given part of a form was a linear function of the probability that that part had permitted correct discrimination in his past. The implications for differentiation theory are discussed.

Several experimental investigations indicate that in discrimination tasks, particular parts of a contour may be more critical than the characteristics of the whole contour. For example, French (1954) found that dot patterns of 7 elements were more easily identified than those of 3 or 12 elements. French explains this by asserting that when the 'point of optimum detail' is exceeded the subject can no longer perceive the pattern as a whole and can respond only to parts, so that he often fails to discriminate correctly.

Anderson and Leonard, working with histograms (14–16 sides), also infer that subjects respond only to parts of forms, since discrimination was easiest for completely random figures. They assume, from inspecting their figures, that randomness provides distinctive details, and they hypothesize that "the advantage of random figures in a recognition task must result from [the subject's] ability to make his discrimination on the basis of distinctive parts of the figure" (1958, p. 269). Likewise, Rappaport, in a study of the role of pattern redundancy, found that his subjects (verbal reports) "almost unanimously agreed that . . . they made little or no effort to look at the whole figure but examined instead some part of the figure that was sufficient to identify it. If one detail was not sufficient for positive identification . . . , then . . . they examined two or three details" (1957, p. 6).

This assumed response to parts of the contour cannot be directly measured in the design of any of these experiments, and hence such inferences must be regarded as speculative. Attneave, on the other hand, used forms constructed by successively mirroring one part, to measure more directly the assumed tendency of subjects to respond to limited portions of a pattern's contour. He found only slight evidence for part perception, as his subjects "made only limited use of the potentialities for economical encoding inherent in symmetrical figures"—and they did so in only one of three experiments (1955, pp. 213–214).

In attempting to determine the role of certain parts in form discrimination, the complexity of the form should be carefully considered. The human face alone, for instance, consists of more than half a hundred lines and arcs, and it seems unlikely that all or most of so complex a form's elements would have important roles in discrimination. The experimental literature of form perception deals with forms much simpler than those encountered outside the laboratory: seldom has the number of elements exceeded two score (40), and often it has been less than a dozen. The simplicity of the experimental material may well account for the scant evidence for discrimination on the basis of distinctive parts, and an underestimation of its importance. Even Gibson (1969), who has constructed an extensive theory of form discrimination on the basis of distinctive features, presents this type of discriminative response to part of a contour as applicable to perception of forms by children but difficult to demonstrate in the case of adults. It would seem, however, that in the case of forms which have a higher, more realistic level of complexity, it is reasonable to hypothesize that adults do respond perceptually primarily to that part of a form's contour which has been most often different from comparison forms; when a form is relatively simple, on the other hand, their perceptual response may be based equally on all portions of its contour. Experiment I was designed to test this hypothesis.

EXPERIMENT I

'Perceptual response' for the purpose of this experiment is defined as discrimination, following the suggestion of Gibson and Gibson (1955).

Method

—Stimulus construction—Two sets of 72 forms, a complex set and a simple set, were constructed to test the hypotheses. For the complex set, forms consisting of 48 line segments were generated on a 6-in. grid ruled in 1/4-in. gradations so that their contours were divided into eight regions of equal area containing 6 line segments each (Figure 1).

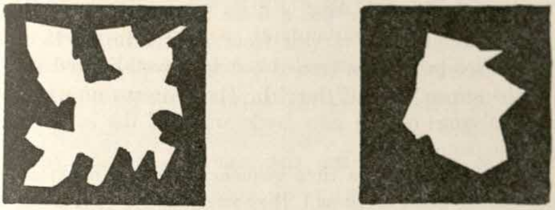


Fig. 1. Standard patterns for complex set and for simple set, Experiment I

The first form generated was designated the *standard*. The next 63 forms generated were *training-comparison* forms. These training forms differed from the standard in the arrangement of the line segments in two regions but were identical to the standard in the other six regions. The comparison forms were divided into eight subsets of 21 members each, every form being used in more than one subset. The members of each subset shared the characteristic of differing from the standard in one region, designated the 100% region. The other seven regions were designated 14% regions, since one-seventh of the forms differed from the standard in each of the regions. Thus, subset 1 consisted of three comparison forms differing from the standard in regions 1 and 2, three differing in regions 1 and 3, and so on. The balanced method of assignment of regions to the 100% and 14% conditions prevents the probability factor from being confounded by differences between regions, since each region acts as its own control.

Eight more forms, designated *test-comparison* forms, were generated, each differing from the standard in a single, different region, so that each region was represented by one of the test-comparison forms.

The set of 72 simple forms (1 standard, 63 training-comparison forms, and 8 test-comparison forms) was generated in the same way, except that there were only 2 line segments per region, so that each form's contour consisted of 16 line segments.

The patterns were cut out, photographed on a black background, and made into 2 by 2 in. negative slides.

—Subjects—The subjects were 160 male and female volunteers from an introductory psychology course. They received extra credit in the course for participating in the experiment.

—Apparatus—Slides were projected on a 9 ft by 8 ft 6 in. screen by a Kodak RA950 Carousel projector. Time of presentation was controlled by a Wollensak tachistoscopic shutter connected in series with a Hunter 111D timer, a Lafayette interval timer, and a General Radio Variac.

—Procedure—The subjects were given a response sheet with these instructions on it: "This experiment is designed to show how people recognize patterns. You are going to see randomly drawn white polygons with many irregular sides. One polygon will be the standard. It will come on for one second, then there will be a five second pause. Next, another polygon called the comparison will come on for one second. Your task is to tell whether the polygons are identical or different. You write 'I' on your paper if they are identical, and 'D' if they are different." There was a 5-sec pause for responding. The subjects were seated at distances varying from 11 ft 2 in. to 18 ft 10 in. from the screen. The experiment was carried out in a well-lighted room; ambient luminance of the screen was 1.07 log ftL. The patterns appeared to the subjects as white polygons on the gray background of the screen and were approximately 60 in. square.

The subjects were shown the time sequencing, using a cartoon figure and a blank slide. During this experiment, they were always told when the patterns were about to appear on the screen.

—Training—Eight groups of 10 naive subjects each were trained on the eight subsets of complex training-comparison forms, one group on each subset. The groups were trained separately in sessions lasting approximately 45 min. Each group was required to discriminate between 42 pairs of forms. The first member of each pair was the standard. The second member of 21 of the pairs was a replica of the standard, which was thus used as a comparison form. The second members of the other 21 pairs were the seven training-comparison forms from the particular subset on which the group of subjects was being trained, which were thus presented three times each. Presentation of the pairs of forms was randomized, with the restriction that none of the seven comparison forms appear twice before each of the others appeared once and that none appear three times before the others had appeared twice, so that the training data can conveniently be analyzed in terms of three 'stages' of training. Subjects were not apprised of these stages. This training establishes the 100% region as a distinctive feature that provides a sufficient basis for correct discrimination.

Eight more groups of 10 naive subjects each were similarly trained on the simple forms. Groups trained on simple forms were run alternately with groups trained on complex patterns.

—Testing—At the end of their respective training sessions, the eight experimental groups trained on complex subsets of patterns were tested with the eight complex test-comparison forms. These forms differ from the complex standard in only one region (one differing in region 1, one differing in region 2, and so on). The test consisted of one trial on each of 16 pairs of forms. The first member of each pair was the standard. The second member of 8 of the pairs was a replication of the standard, used as a comparison form. The second members of the other 8 pairs were the eight test-comparison forms. The 16 pairs were presented in random order. Subjects were not informed that training had ended and testing begun, nor of the number of test-comparison forms or standard replications. The experimental hypothesis predicted that subjects should make relatively fewer errors discriminating the standard from test-comparison patterns differing in the 100% regions than in the 14% regions.

Results

Lindquist's Type I analyses of variance (1953, p. 178) were employed throughout, with the group's scores, not the individual subject's scores, as the unit of analysis. Use of the group's scores provides a very conservative analysis but is particularly necessary for the test data, since each subject's score per cell consists of a 0 or 1 depending on whether or not he erred on the test-comparison form for a given contour region.

The improvement during training ($F = 12.76$, $df = 2, 28$, $p < .01$) is shown in Figure 2. Although the overall effect of complexity was not significant ($F = 1.95$, $df = 1, 14$, $p > .05$), a high level of performance was reached more quickly on the complex forms, resulting in a significant interaction of complexity and training ($F = 7.27$, $df = 2, 28$, $p < .01$). This result is similar to the one

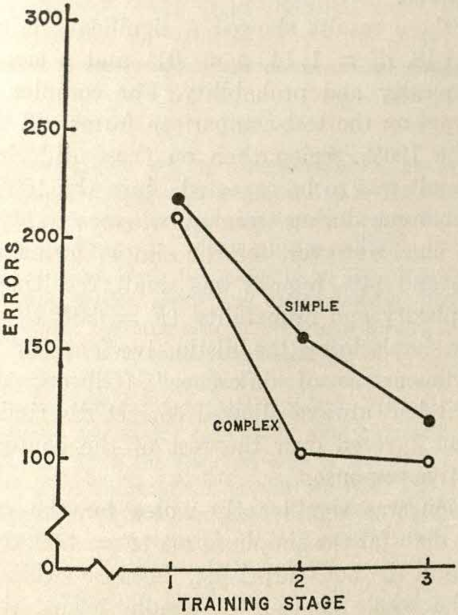


Fig. 2. Total error scores during training on complex and simple patterns, Experiment I

that Deese obtained for his regular polygons but not for his irregular ones (1956). He was led to suggest that the processes of discrimination might depend on the type of polygon. However, since the forms used in the present experiment were certainly 'irregular' rather than 'regular' in the context of Deese's experiment, the 'type' of form is probably not the reason for the difference in results he encountered. Instead, the level of complexity seems crucial, since the forms used in this experiment were a good match for Deese's regular polygons in terms of complexity, while his irregular polygons were all much simpler.

The purpose of the testing procedure was to determine whether the subjects responded to all portions of the form's contour. Accordingly, test-comparison forms were presented that differed from the standard in one or another portion of the contour. Failure to perceive a particular portion of the contour would result in a false report that the respective test-comparison form was identical to the standard. These mean failures to detect a difference in a region of contour were 3.13 and 3.48 for the 100% and 14% regions of simple patterns, and 3.30 and 4.74 for the 100% and 14% regions of complex patterns.

Analysis of these results showed a significant effect of probability ($F = 11.18$, $df = 1, 14$, $p < .01$) and a significant interaction of complexity and probability. For complex forms, there were fewer errors on the test-comparison forms differing from the standard in the 100% region than on those differing in a 14% region. This result was to be expected, since the 100% region was always a reliable cue during training, whereas a 14% region was not a reliable cue. However, for the simple forms the difference between 100% and 14% regions was small, resulting in an interaction of complexity and probability ($F = 6.35$, $df = 1, 14$, $p <$ .05). Thus, for simple forms the 'distinctive features' that had provided real "dimensions of difference" (Gibson, 1969)—regions of contour that had always allowed correct discrimination of the form—were not favored over the rest of the contour as a basis for discriminative response.

Discrimination was significantly worse for the complex forms during testing than for the simple forms ($F = 4.69$, $df = 1, 14$, $p <$ .05). This result is not surprising, because while subjects responded to the whole contour of simple forms, they showed a tendency to base their response on one part of the contour of complex forms. During testing, this part of the contour—the

100% region—was not a reliable cue, although it had been an excellent cue during training. Responding to it produced extra errors during testing, if the form was complex, since the rest of the contour was likely to be ignored.

EXPERIMENT II

The hypothesis that subjects would respond more to some parts of the contour of complex forms than to other parts was substantiated and, as predicted, the 'distinctive features' concept had less validity when applied to simple forms. The subject apparently responds to all of a simple form's contour with little extra emphasis on parts particularly useful in discrimination. However, the direction of the difference in number of errors for the 100% and 14% regions was the same. This may indicate that part of a subject's perceptual response to a form is based on characteristics besides its most distinctive features. Experiment II was conducted to determine how reliable a cue a portion of contour has to be in order for a subject to base his discriminative response on it. To answer this question a range of probabilities was tested with complex forms.

Method

—Stimulus construction—A set of complex forms having 48 sides was generated by the method used in Experiment I. The first form generated was designated the *standard*. Eight subsets of *training-comparison* forms were also generated, each subset consisting of nine forms. All nine training-comparison forms in any set differed from the standard in the arrangement of the six line segments in one region, which was designated the 100% region. Six of these same comparison patterns also differed in a second region, designated the 67% region. Three of the nine comparison forms also differed in a third region, the 33% region. For all nine forms, a fourth region was designated the 0% region and never differed from the standard. The interactions of these different contingencies results, for each set of nine training-comparison forms, in two forms differing from the standard only in the 100% and 67% regions, one form differing from the standard only in the 100% and 33% regions, and two forms differing from the standard only in the 100%, 67%, and 33% regions. Each region acts as its own control, appearing equally often as a 100%, 67%, 33%, and 0% region. Therefore, significant differences between the probabilities cannot be confounded by differences between regions.

Next, 16 *test-comparison* forms were generated, each form differing from the standard in only one region. Thus, there were two test-comparison forms differing from the standard in region 1, two differing in region 2, and so on.

—Subjects—Eight groups of 10 new subjects each were obtained from introductory courses for this experiment.

—Procedure—The procedure was identical to that of Experiment I, except that there were 54 training trials instead of 42, since there were more training forms. Eight groups of 10 subjects each were run, one group on each of the eight sets of training-comparison forms.

Results

The improvement during training was linear and significant ($F = 63.71$, $df = 1, 14$, $p < .001$). When the subjects were tested on these complex forms, probability had a significant effect ($F = 4.96$, $df = 1, 18$, $p < .05$), corroborating Experiment I. The unit of analysis was again the group totals for failures to discriminate the standard from the comparison form. The effect was linear; nearly twice as many errors were made on the 0% region, which had never been a dimension of difference during training, as were made on the 100% region, which was always a dimension of difference during training (note Figure 3). All comparisons of individual treatment means were significant at the .01 level according to Duncan's multiple-range test, except between the 67% and 33% means. Although the total number of errors was not great, some parts of a form's contour demonstrably played a more important role in discrimination of the form than other parts. Although fewer errors were made on 'corner' regions at all probabilities, this position effect (corner versus medial regions) fell short of significance ($F = 5.30$, $df = 1, 5$, $p > .05$), and there was no significant interaction of probability and position ($F = .44$, $df = 3, 18$, $p > .05$).

DISCUSSION AND CONCLUSIONS

The major implications of Experiments I and II are quite clear. Subjects have a tendency to respond to those parts of a complex form's contour that have in the past differentiated this form from others. There is a linear relation between the degree of this response emphasis on a particular part and the past probability of differentiation by this part. However, for simple forms the subject apparently responds to all of the contour equally; there is little response bias for different parts of the contour, regardless of differences in the extent to which they have differentiated between forms.

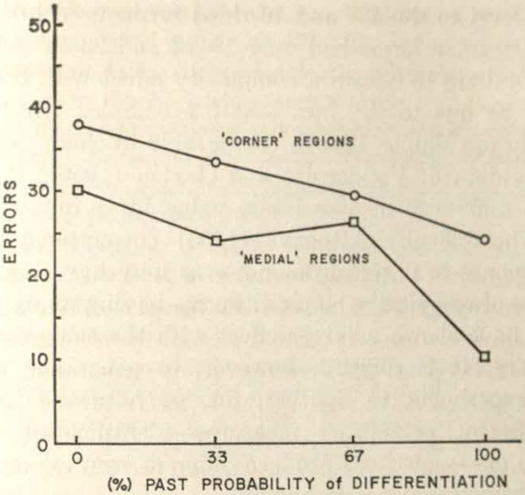


Fig. 3. Total failures to discriminate the standard from the comparison form, as a function of probability for 'corner' and 'medial' regions, Experiment II

The results of Experiment I also indicate that interpreting part perception in terms of the 7-unit 'span of apprehension' may be premature, since perception of the whole contour persisted in patterns of 16 sides (French, 1954). The experiment by Deese indicates that subjects may identify forms of this complexity level by relating them to some familiar object. In his words, "it seems reasonable to suppose that in the case of the forms that are coded they are perceived as a whole and responded to on the basis of a general impression of their shape. Forms that are infrequently coded are probably identified on the basis of some salient features of the contour and thus identified by part" (1956, 56–60). Although Deese used a longer—10 sec—exposure time and different types of forms, his supposition is supported by the direct evidence of Experiment I that discrimination of complex forms is more likely to be based on specific parts, while for simple forms all parts of the contour have an equal role in discrimination.

The relevance of the finding that a high level of complexity was necessary for part perception to become important may be inferred from the fact that Deese obtained the expected differences between responses to simple and complex 'regular' forms but not 'irregular' ones. However, while his complex regular forms (45–55 sides) and simple regular forms (16–17 sides) are comparable in complexity level to the 48- and 16-sided forms used in the present study, his irregular forms had only 10–12 and 23–27 sides respectively. The failure to obtain a complexity effect with the irregular forms may be due to the fact that the complex irregular forms are probably too simple, though comparable to those used in other studies. The data of Vanderplas and Garvin (1959), for instance, indicate no difference in association value for forms with 12 and 24 sides. The difficulty Attneave (1955) encountered in demonstrating response to parts of dot patterns may have been also due to the levels of complexity he used. Forms having many more sides might well have shown a larger effect with the same experimental manipulations. It is difficult, however, to generalize confidently from one experiment to another, for, as Attneave has demonstrated, different perceptual measures—identification, reproduction, and so on—sometimes produce different results, and time and stimulus parameters also vary widely.

Since the forms with which subjects ordinarily deal in the nonlaboratory world are much more complex than those used in experimental research, it seems appropriate to conclude that experience may often have an important role in determining what features of these complex contours are the basis for discriminative response. Those parts that have consistently marked a difference between forms are favored over those that have only sometimes or rarely denoted a difference, except in the case of simple forms. Differentiation theory lends itself to this interpretation of the present data (see Gibson, 1969) and could profitably be modified to account for the effect of complexity and to define more precisely the role of experience.

Notes

This research was supported in part by USPHS Grant MH-1061-1 from the National Institute of Mental Health. The author thanks R. B. Freeman, Jr., for his help in formulating and conducting the experiment, and James Starling for suggesting methods of statistical analysis. Received for publication January 20, 1969.

References

Anderson, N. S., and Leonard, J. A. 1958. The recognition, naming and reconstruction of visual figures as a function of contour redundancy. *Journal of Experimental Psychology* 56:262–270.

Attneave, F. 1955. Symmetry, information, and memory for patterns. *American Journal of Psychology* 68:209–222.

Deese, J. 1956. Complexity of contour in the recognition of visual form. U. S. Air Force WADC technical report 56–60.

French, R. S. 1954. Identification of dot patterns from memory as a function of complexity. *Journal of Experimental Psychology* 47:22–26.

Gibson, E. J. 1969. *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Gibson, J. J., and Gibson, E. J. 1955. Perceptual learning: Differentiation or enrichment? *Psychological Review* 62:32–41.

Lindquist, E. F. 1953. *Design and analysis of experiments in education.* Boston: Houghton Mifflin.

Rappaport, M. 1957. The role of redundancy in the discrimination of visual form. *Journal of Experimental Psychology* 53:3–10.

Vanderplas, J. M., and Garvin, E. A. 1959. The association value of random shapes. *Journal of Experimental Psychology* 57:147–154.

IMAGERY IN ASSOCIATIVE LEARNING AND MEMORY

Martin Bolt
Michigan State University

Three experiments were conducted to determine under what conditions, if any, presenting picture cues on the learning trial can facilitate the free recall of nouns. The presentation of picture cues facilitated the free recall of concrete nouns (Experiment I) but not of abstract nouns (Experiment II). The influence of the arrangement of pictures cues was unclear, since arrangement influenced the number of cues the subjects could recall (Experiment III) but not the number of response words (Experiment I).

Mnemonic systems that involve the use of imagery and the formation of associations between a well-learned series of words ('peg' list) and the set of words to be remembered have recently undergone experimental investigation (see Bugelski, Kidd, and Segmen, 1968; Paivio, 1968; Wood, 1967a). Verbal cues have been used as pegs, and the subjects have usually been instructed to transform the verbal cues to images of objects before employing them. The present study attempted to extend the previous work on mnemonic devices by investigating the effectiveness of pictures as pegs.

Earlier studies investigating the use of verbal cues in free recall have indicated that the presence of a cue during learning has no effect unless the cue is also presented at recall (Crouse, 1968; Tulving and Osler, 1968; Wood, 1967b). However, until the present experiment there had been no attempt to investigate whether presenting a picture cue on only the learning trial would facilitate recall if the subjects were instructed to associate the to-be-remembered words with the pictures. It seemed plausible that the test-trial pictures, being easier to recall than the words (see Paivio, Rogers, and Smythe, 1968), might thus 'mediate' recall of the to-be-remembered list. The arrangement of the pictures was also manipulated, since it was felt that the 'presentation order' of the pictures might influence the availability of the cues and, consequently, the extent to which the pictures facilitate the recall of the words to be remembered. The arrangement of the picture cues was expected to influence performance when the picture cues were presented for the study trial only, but not when the picture cues were also presented for the recall trial. Word cues were expected to facilitate recall only when presented on both learning and recall trials.

EXPERIMENT I

Method

—Design and materials—A 3 × 2 factorial was used, with the nature of cues (organized pictures, unorganized pictures, or words) and the nature of recall (cues or no cues) being manipulated. An additional control group received no cues and was given standard free-recall instructions.

Each of the three sets of cues consisted of 32 items presented on a 12 by 18 in. sheet of paper. The organized-pictures set was composed of 32 colored pictures depicting a highway intersection (see Reynolds, 1966) that included a gas station, shopping center, farm, church, truck, bus, mountains, trees, and so on. The unorganized-pictures set consisted of the identical pictures presented in a random arrangement. The third set of cues consisted of 32 words in the same positions that the 32 organized pictures occupied. The word cues were the labels for the 32 pictures.

A response (to-be-remembered) list of 32 words was used. The response words were selected from the Paivio, Yuille, and Madigan norms (1968) and had concreteness values ranging from 6.38 to 7.00 with a mean rating of 6.86

—Procedure—The subjects were 98 students from introductory psychology courses at Michigan State University. Fourteen were randomly assigned to each of seven groups. All subjects were run individually and were given such general information about the experiment as the nature and number of words to be recalled and the rate at which the words would be presented. The six groups receiving cues were given instructions on their use. That is, the subjects were told to associate the response words with the cues through imagery. They were told they would not have to recall the cues; they were also told whether the cues would be available at recall.

The response words were read at a 5-sec rate. Following the presentation of the response list, the subjects had as much time as they wanted for written recall. No subject took longer than 5 min for recall.

Results and discussion

The mean numbers of words correctly recalled for the six experimental groups are presented in Table 1. Picture groups were

Table 1. Mean scores for the six experimental groups of Experiment I

Nature of recall	Nature of cues: Words	Unorganized pictures	Organized pictures
No cues			
M	13.36	17.50	18.36
SD	3.64	3.50	3.39
Cues			
M	17.79	19.57	18.93
SD	5.60	4.27	2.69

superior to the corresponding word groups, and groups presented cues at both learning and recall tended to do better than groups presented the cues only at learning. A 3 × 2 analysis of variance indicated that both the nature of the cues [$F(2, 78) = 5.06$, $p < .01$] and the presence of cues during recall [$F(1, 78) = 6.93$, $p < .05$] were significant. However, neither the interaction between the main effects [$F(2, 78) = 1.57$, $p > .05$] nor an orthogonal comparison between the organized- and unorganized-pictures conditions was significant ($F < 1$).

An orthogonal comparison between the average of the picture conditions and word condition was significant [$F(1, 78) = 10.10$, $p < .01$]. Since previous research has indicated a superiority of pictures over words both in free-recall and paired-associate learning, it seems plausible that the superiority of picture cues was due to the better recall of the picture cues and the greater effectiveness of pictures as cues than words as cues.

The mean number of correct responses for the control group was 13.14. Although there was no significant difference between the control group and the experimental group presented word cues only during learning ($F < 1$), the performance of the unorganized-pictures group presented cues only at learning was significantly better than that of the control group [$F(1, 91) = 7.76$, $p < .01$]. Since the mean scores for the four remaining experimental groups were higher than that of the unorganized-pictures group presented cues only at learning, the subjects in the four remaining groups were significantly better than the control subjects.

Thus, the results of Experiment I indicate that picture cues presented only on the learning trial can facilitate recall but that word cues presented only during learning do not facilitate recall. The purpose of Experiment II was to determine the reliability and generality of this finding by investigating whether presenting cues only during learning can facilitate the recall of abstract as well as concrete words.

EXPERIMENT II

Method

—Design and materials—A 2 × 2 factorial was used in which cuing (picture cues or no cues) and the nature of the response list (concrete or abstract words) were manipulated. The 32 organized pictures used in Experiment I were the only cues used in this experiment. Two response lists of 32 words each were selected from the Paivio, Yuille, and Madigan norms (1968). The mean concreteness value was 6.82 for the concrete list and 1.69 for the abstract list.

—Procedure—The procedure was essentially identical to that of Experiment I except that the nature of recall was not manipulated. All subjects who received cues received them only during the learning trial. The 56 students from introductory psychology courses at Michigan State University who participated in this experiment were randomly assigned to the four conditions such that 14 subjects were in each condition.

Results and discussion

The mean numbers of words correctly recalled were 19.57, 15.21, 12.64, and 13.50 for the cued-concrete, concrete-control, cued-abstract, and abstract-control groups respectively. An analysis of variance on these data indicated that the nature of the response list significantly influenced performance [$F(1, 52) = 6.08$, $p < .05$] but that cuing did not [$F(1, 52) = 2.74$, $p > .05$]. The most important result of this experiment was the significant interaction between response list and cuing [$F(1, 52) = 6.07$, $p < .05$]. This finding indicates that the effectiveness of picture cues in facilitating recall depends on the nature of the items being recalled. Although picture cues presented during learning can improve the recall of concrete nouns, picture cues are not effective in facilitating the single-trial free recall of abstract nouns. The subjects who received the abstract list probably had difficulty in forming associations between the pictures and the abstract words. Perhaps with additional practice they would have been able to establish these associations and thus use the picture cues effectively.

EXPERIMENT III

An unexpected finding of Experiment I was that unorganized pictures were as effective as organized pictures in mediating recall of the response list when the cues were presented only at learning. It had been predicted that if picture cues were presented only at learning, then the recall of groups presented the pictures in a meaningful arrangement should be superior to the recall of groups presented the pictures in a random arrangement. The failure to find higher recall for the subjects who received the meaningful arrangement suggests that the arrangement did not influence the number of picture cues recalled or that the number of picture cues did not influence the number of response words recalled. The purpose of Experiment III was to test whether the arrangement of picture cues influences the number of cues recalled.

Method

One group was presented 32 pictures in a meaningful arrangement, and a second group received the same pictures in a random arrangement. A third group was presented the 32 word labels of the 32 pictures. The word cues were presented in the same positions that the pictures occupied in the organized-pictures group.

The materials utilized were the three sets of cues from Experiment I. The only difference was that in this experiment the names of the 32 pictures were printed next to the appropriate objects. The subjects were 48 students from introductory psychology courses at Michigan State University; 16 subjects were randomly assigned to each of the three groups. All subjects were run individually and were given standard free-recall instructions. The pictures or words were presented for 30 sec. Following presentation each subject was instructed to write down as many words, or in the case of pictures, as many labels, as he could recall.

Results and discussion

The mean numbers of correct responses were 15.19, 11.44, and 9.00 for the organized-pictures, unorganized-pictures, and word groups respectively. An orthogonal comparison between the organized- and unorganized-pictures conditions was significant [$F(1, 45) = 13.04$, $p < .01$], indicating that pictures presented in a meaningful arrangement are more easily recalled than pictures presented in a random arrangement. An orthogonal comparison between the average of the picture groups and the word group was also significant

$[F(1, 45) = 22.98, p < .01]$, indicating that pictures are more easily recalled than words. This is consistent with earlier studies (e.g., Paivio, Rogers, and Smythe, 1968) that have found recall to be higher for pictures than for their labels.

The superior recall of 'meaningfully' arranged pictures over randomly arranged pictures is consistent with earlier studies. Earlier studies, however, have used verbal items as materials, with organization being manipulated through presentation order (e.g., Earhard, 1967; Tulving, 1965). The present experiment indicates that the relationship between presentation order and recall maintains even when pictures are used instead of words.

In light of the findings of Experiment I, the fact that meaningfully arranged pictures were more easily recalled than randomly arranged pictures suggests that the availability of additional cues does not necessarily result in better response-word recall. If the results of Experiment III accurately indicate the number of cues that were available to the subjects in Experiment I, it appears that there is little difference in recall of response words when the subjects are able to recall 11 cues (unorganized pictures) than when they are able to recall 15 cues (organized pictures).

Notes

This paper is based on a doctoral dissertation submitted to the Department of Psychology, Michigan State University. The invaluable assistance and suggestions of Gordon Wood, chairman of the author's committee, are gratefully acknowledged. Thanks are also due Joe L. Byers, Donald M. Johnson, and James L. Phillips. The author is now at Calvin College, Grand Rapids, Michigan. Received for publication August 11, 1970.

References

Bugelski, B. R., Kidd, E., and Segmen, J. 1968. Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology* 76:69–73.

Crouse, J. H. 1968. Storage and retrieval of words in free-recall learning. *Journal of Educational Psychology* 59:449–451.

Earhard, M. 1967. Subjective organization and list organization as determinants of free-recall and serial-recall memorization. *Journal of Verbal Learning and Verbal Behavior* 6:501–507.

Paivio, A. 1968. Effects of imagery instructions and concreteness of memory pegs in a mnemonic system. *Proceedings of the 76th Annual Convention of the American Psychological Association* 3:77–78.

Paivio, A., Rogers, T. B., and Smythe, P. C. 1968. Why are pictures easier to recall than words? *Psychonomic Science* 11:137–138.

Paivio, A., Yuille, J. C., and Madigan, S. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology* 76(1, pt. 2).

Reynolds, J. H. 1966. Cognitive transfer in verbal learning. *Journal of Educational Psychology* 57:382–388.

Tulving, E. 1965. The effect of order of presentation on learning of "unrelated" words. *Psychonomic Science* 3:337–338.

Tulving, E., and Osler, S. 1968. Effectiveness of retrieval cues in memory for words. *Journal of Experimental Psychology* 77:593–601.

Wood, G. 1967a. Mnemonic systems in recall. *Journal of Educational Psychology* 58(6, pt. 2).

Wood, G. 1967b. Category names as cues for the recall of category instances. *Psychonomic Science* 9:323–324.

SEMANTIC ORGANIZATIONS AND MEMORY FOR RELATED SENTENCES

Ralph Y. Sasson
University of Alberta, Canada

Two experiments investigated the effect, on recall, of various organizations of related sentences, organizations based on meaningful relationships between the sentence parts. Learning was highest when the organizational sequence of sentences was both thematic and temporal, lowest when it was temporal, and intermediate when it was based on the logical subjects of the sentences. The different parts of the sentences were differently recalled depending on the organization.

Research directed at uncovering the variables operating in the learning and recall of meaningful verbal material has neglected what is thought to be a very important problem: the organizational roles played by the relationships between parts of syntactically correct sentences. This problem is important because meaningful interpart relationships may frequently play a central role in the recall and organization of material. Consider an individual who is studying the French Revolution of 1789. This individual will 'master' the material better if he perceives how one event gave rise to another and attempts to organize the material in a logical and coherent manner than if he tries to blindly memorize the events involved. His memory of the different parts of the material might possibly also be different depending on the organization used when the material was being learned.

By the phrase 'meaningful ways of relating parts of the material' is meant, then, ways that relate the parts according to their semantic features (e.g., cause and effect features, temporal features), and not according to arbitrary incidental characteristics or artificial mnemonic devices.

The notion that meaningful interpart relationships exercise an effect on the organization of meaningful verbal material or sentences (as in this study) is not new. Bousfield (1953) and many other subsequent investigators have demonstrated that subjects in free recall tend to retain words or items in clusters when these items are members of a common conceptual category. Since sentences can also be conceptually related, can pertain to the same theme, and can have other meaningful relationships between each other, it is not unreasonable to expect that sentences having such relationships may be recalled in clusters.

Now, there are many possible ways in which information from related sentences can be organized. Analogously to the category-clustering phenomenon, sentences can be related around a common theme; when the subject is asked to recall the sentences, he first recalls the theme and then 'reconstructs' the sentences. Such a *thematic organization* implies that information from sentences in, say, a historical passage is filed or ordered according to the nature of events (and not according to less important details such as dates or names) and that events pertaining to the same topic are grouped under the same superordinate label (theme). Thus, in order to retrieve other information, the subject would usually have to recall the event first in order to retrieve those other details.

Other organizations, however, are possible. In a *temporal organization* each event and other information can be 'listed' or filed under its particular date of occurrence. The sequence of dates can be temporally ordered; the subject goes through this sequence, recalling at each date the particular event. This could be an efficient plan to retrieve items from memory. The thematic and temporal kinds of memory structures (which is probably too simple a distinction) are schematically presented in Figure 1. Note that in the thematic organization there is no hierarchy (temporal or other) between events. In the temporal organization, however, dates are temporally hierarchized.

One way to investigate how information from related sentences is organized and how different organizations affect memory for sentences is to have different groups of subjects learn the same sentences in different sequences, each sequence being based on or governed by a particular kind of semantic interpart relationship, and then to compare the groups with respect to their learning performance. By using this method, as we did, we in effect made it possible for interpart relationships to provide a basis for organization and then determined whether these different kinds of memory

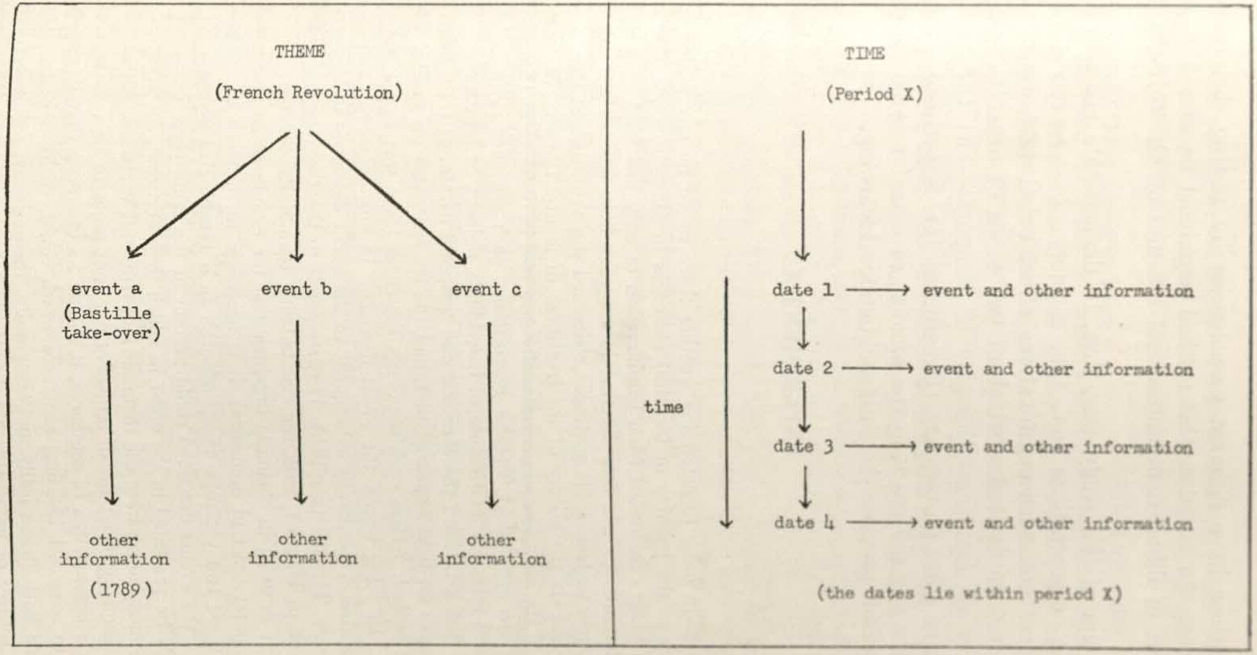

THEME
(French Revolution)
event a
(Bastille take-over)
event b
event c
other information
(1789)
other information
other information
TIME
(Period X)
time
date 1 → event and other information
date 2 → event and other information
date 3 → event and other information
date 4 → event and other information
(the dates lie within period X)

Fig. 1. Hypothetical thematic and temporal memory organizations, or storage patterns, for information from sentences that can be differently related

organizations have different prepotencies by testing their effects on learning. In addition, this method permitted an assessment of the effect of different organizational schemes on the memory for different sentence parts.

The aim of this study was, first, to demonstrate the effect of semantic organizations on the memory for sentences and to discover how such organizations affect recall (Experiment I), and second, to determine which of two kinds of organizations or sequences of sentences—thematic or temporal—has the greater facilitative effect on memory (Experiment II). Experiment II also tested for any difference in the amount of transfer to subsequent learning from previous thematic and temporal learning.

EXPERIMENT I

Method

—Material—In order to determine whether there would be any performance differences in the learning of identical sentences in different sequences, each sequence being constructed from consideration of only one kind of intersentence relationship, three lists differing in sequence were constructed out of a dozen factual sentences. Six sentences concerned the evolution of the concept of light, and six the history of the treatment of mental illness. The veracity of some of the sentences was occasionally distorted in order to fit the requirements of the experiment. Besides *propositions,* each sentence included *dates* or temporal messages, and *names* of individuals, objects, or locations.

In the first list, called the *thematic and temporal within, or CTW, list,* the six sentences on light appeared first, and the six sentences on mental illness appeared second. Within each set of six sentences there was also a temporal sequence. In other words, the event described in any sentence (except the first) was always temporally subsequent to the event described in the preceding sentence. This list (see Table 1) corresponds to the thematic organization represented in Figure 1, with the events being temporally ordered.

While it is usual in thematic organizations for events to be presented in their temporal order of occurrence, it is necessary to measure the effects of each of these two variables independently. In a meaningful passage, for example, all the events described could pertain to the same topic but need not necessarily be presented in the order in which they occurred. An attempt was therefore made to separate the thematic and temporal variables. Although the second (thematic, or C) list of sentences may not seem to be very similar to a thematic organization of sentences, it should be observed that it was quite difficult to construct identical sentences which could, without being bizarre, be ordered both thematically and temporally, thematically, and according to sentence subject or name (the third variable investigated; note that in Experiment II, which did not manipulate the third variable, the thematic organization is identical to a real one).

The sentences in the second list, called the *thematic, or C, list*, were arranged in a thematic sequence only. The six sentences on light appeared first and the six sentences on mental illness appeared second; however, within each theme the events were not ordered in temporal sequence. The first four sentences of the C list appear in Table 1.

The third list was called the *sentence subject (name, or N) list*. Here the sentence that was chosen to follow another one in a given pair was the sentence in which was embedded the name of the same individual (or object or location) as in the preceding sentence. Since each name was mentioned in two different sentences, once in a sentence on light and once in a sentence on mental illness, there was in this list a constant alternation of themes from one sentence to the next. This prevented the sentences from being thematically ordered, and since each name had two different dates associated with it, it also prevented the sentences from being temporally ordered. The aim behind presenting such a list was a desire to see whether the subjects in the experiment would use these names as a basis for organization. The first four sentences in the N list are also shown in Table 1.

—Design and procedure—Three independent groups listened to one of the three lists of sentences as the list was played twice on a tape recorder. Immediately after the second presentation of the list, the groups were given five serial anticipation-learning trials on the CTW list.

Fifty-eight subjects, 31 female and 27 male, all enrolled in the introductory psychology course at the University of Alberta, were assigned randomly to the three groups. Groups CTW and C contained 20 subjects each, and group N contained 18. The subject was seated at a table facing a 50 by 50 in. screen with his back to the experimenter, who was seated at another table. This was done in order that the subject not see any unintentional responses of the experimenter (e.g., facial expressions) which might influence his performance. A Kodak AV-900 Carousel slide projector, by being connected to two Hunter 111 timers (which were themselves interconnected by appropriate wiring), worked automatically. Positive (black on white) slides were made of each sentence. The slides were placed in the projector so that between successive slides the screen was illuminated by the projector light. This provided a recall period.

Immediately after the second (aural) presentation of the list, the subject attempted to recall the sentence he had heard first. When recall time (30 sec) was over, the first sentence was flashed on the screen and remained exposed for 13 sec. Immediately after this sentence was flashed off, the subject attempted to recall the sentence he had heard second; after recall time elapsed he was shown that sentence, and so on. This procedure was continued for five trials, a trial consisting of going through the entire list (all sentences) once. No exchanges of any sort occurred between subject and experimenter during the recall session; all that the experimenter said was "Next" after a sentence was automatically flashed off. Intertrial interval was zero.

—Analysis of the sentences and scoring—The sentences were divided into parts according to dates or times mentioned (e.g., time immemorial, earlier); names of individuals, locations, or objects (e.g., Witch Hammer); and propositions. A proposition was defined, in consonance with tradition, as a declarative sen-

Table 1. Examples of sentences used, Experiment I

The CTW list

1. Ever since time immemorial/man has wondered about the nature of light./
2. In 1704,/Newton/in England/thought that light is made up of particles./
3. In 1750,/ it is interesting to note,/Lebon/said if light consists of particles,/ why do we not feel the particles bouncing on our eyes./
4. Huygens,/in the late 18th century,/was unable to answer the scientist's question,/but conceived light as consisting of waves;/this view urged scientists to do their experiments./
5. In 1803,/Young/proved that light consists of waves/when he passed a beam through two pinholes/and this set up interference patterns/like the waves of water./
6. Later,/in 1840,/the kind of waves that light is made up of was discovered by/ Fresnel/and Arago./
7. In medieval times,/the treatment of demented persons was not one of compassion/and sympathetic listening/but one of brutality/and persecution./
8. Before the publication of the book/the Witch Hammer,/Newton/in 1720/ thought that mental disease was due to being perverse,/witch-natured,/ and possessed of the devil./
9. This atmosphere continued to persist;/in 1740/Huygens/in Germany,/ published a book entitled/the Witch Hammer,/which associated mental disorders with witchcraft/and asserted that the insane should be tortured/ and burnt./
10. In 1780,/Young/questioned the witchcraft theory/by saying if the insane are all witches/why do the insane frequently perform witchcraft unsuccessfully;/ this was the beginning of the attack on the witchcraft theory./
11. The first step forward was made/in 1793/by Fresnel/and Arago;/the latter freed the insane/from their chains/and encouraged a sympathetic hearing of their troubles./
12. Later,/in 1810,/Lebon/continued the work of/Fresnel/and Arago/and fought for the establishment of mental hospitals./

The C list

1. Ever since time immemorial/man has wondered about the nature of light./
5. In 1803,/Young/proved that light consists of waves/when he passed a beam through two pinholes/and this set up interference patterns/like the waves of water./
2. In 1704,/Newton/in England/thought that light is made up of particles./
6. Later,/in 1840,/the kind of waves that light is made up of was discovered by Fresnel/and Arago./

The N list

5. In 1803,/Young/proved that light consists of waves/when he passed a beam through two pinholes/and this set up interference patterns/like the waves of water./
10. In 1780,/Young/questioned the witchcraft theory/by saying if the insane are all witches/why do the insane frequently perform witchcraft unsuccessfully;/this was the beginning of the attack on the witchcraft theory./
2. In 1704,/Newton/in England/thought that light is made up of particles./
8. Before the publication of the book/the Witch Hammer,/Newton/in 1720/ thought that mental disease was due to being perverse,/witch-natured,/and possessed of the devil./

Note: The word(s) between slashes correspond to the parts, or segments, into which the sentences were divided for analysis.

tence in which anything whatsoever is affirmed or denied. It can also be conceived that the sentences were partly analyzed into 'kernel sentences,' but not always of the active, affirmative form to which Miller (1962) refers when he uses the term.

There were several reasons for dividing the sentences in the manner described. Facts such as dates and names should be given credit if recalled correctly. Furthermore, it was unrealistic to expect, with the type of material used, that the subject would recall it or attempt to recall it in predominantly verbatim fashion. Inevitably the subjects would recall the sentences in their own words (except for dates and names) and in different grammatical structures—while keeping the core meaning (the events) of the sentences unaltered. Hence, the reason for partly dividing the sentences into propositions. The same reason was the basis for allowing within a segment (*a*) the occurrence of active and passive grammatical transformations, (*b*) the substitution of synonyms, (*c*) the permuting of words, and within a sentence (*d*) the permuting of segments (e.g., *Young, in 1803* instead of *in 1803, Young* in sentence 5).

The scoring procedure was based on the subject's recalling the content within the slashes (the part, or segment) correctly. As the subject spoke, the experimenter (who had five copies of the sentences in front of him, one copy for each trial) marked a check if a part within a sentence was correctly reproduced according to the above-mentioned criteria. No partial credit was given if any deviation from the criteria occurred.

Results

Recall of the different types of sentence parts was taken into consideration in the analysis of the data. The arc-sine transformation was also applied to the data in order to correct for heterogeneity of variance. The means (combined over subjects and trials) for percent correct recall were 41.05 for group CTW, 27.90 for group C, and 24.35 for group N. For this data a Duncan's new multiple-range test applied to unequal *N*s was made to determine the significance of the difference between the various means (Kramer, 1956). The results of this test show (as does Figure 2) that recall of the CTW organization was significantly superior to recall of the C organization ($p < .01$) or of the N organization ($p < .01$). Comparison of the C and the N organizations, on the other hand, showed no significant difference ($p > .05$).

Mean recall of dates, names, and propositions, for each organization condition and combined over organizations, is given in Table 2. Analysis of variance for these data showed significant differences between the various means ($p < .01$). Using the error term from this analysis of variance, *t* tests were made to determine the

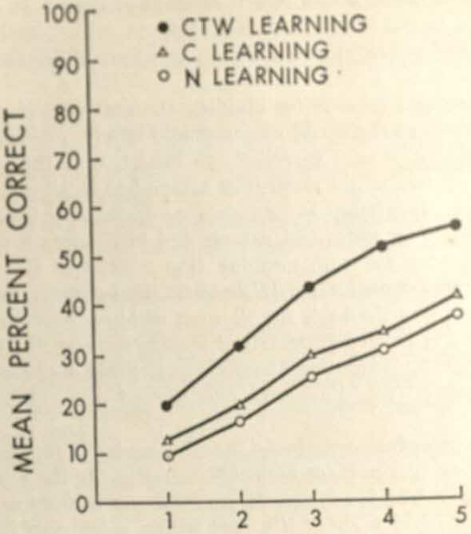


Fig. 2. Mean percent correct recall of segments as a function of trials for the three experimental groups, Experiment I

significance of the difference between means of theoretical interest. Briefly, date recall was higher in the CTW condition than in the C condition ($p < .01$) or the N condition. Name recall was not particularly favored by the N condition, either over the C condition ($p > .05$) or the CTW condition; but proposition recall was favored by the C condition over the N condition ($p < .05$) and was also high in the CTW condition. It was also found that name recall was significantly better than date recall, but not proposition recall, in the N condition ($p < .05$). The overall difference between name recall and proposition recall for all conditions was not significant ($p > .10$).

Table 2. Mean percent correct recall of different sentence segments, Experiment I

	Organization condition			Combined means
	CTW	C	N	
Date recall	19.03	3.32	3.33	8.56
Name recall	6.02	4.52	11.70	7.41
Proposition recall	16.00	17.06	9.32	14.13

EXPERIMENT II

Method

—Materials—Two differently organized lists were constructed out of exactly the same set of sentences used in Experiment I. In the *thematic, or C, list,* the first four sentences dealt with the concept of light and the last four with the history of mental illness. Within each theme no temporal sequence was present, but there was a thematic form of organization: the event described in a sentence was, according to the text, thematically related to another event mentioned in the same theme (see Table 3). Note that unlike Experiment I, however, no list here in Experiment II had more than eight items.

The *temporal, or T, list* was organized on a temporal basis: the event described in a sentence was temporally subsequent to the one described in the preceding sentence. Furthermore, the sentences of each theme alternated in their order of occurrence and did not themselves follow a thematic sequence (see Table 3).

—Design—Two different groups were used. Group CT (18 subjects) underwent four serial-anticipation learning trials on the C list (on each trial they had to anticipate the next sentence) and then four anticipation trials on the T list, each set of trials being preceded by two aural presentations of the C and T lists respectively. Group TC (18 subjects) was given the same tasks but in a reverse order: two presentations of the T list, four anticipation trials on this list, two presentations of the C list, and four anticipation trials on this list.

—Procedure and analysis—The procedure (and the equipment employed) and the analyses were identical to those used in Experiment I.

Table 3. Examples of sentences used, Experiment II

The C list

1. In 1803,/Thomas/Young/because he was skeptical of the idea of/Newton,/ attempted to prove that light consists of waves/by passing a beam through two pinholes/and showing that it set up interference patterns/like the waves of water./
2. Before the experiment of/Young,/in 1704 precisely,/Newton,/having shown that light travels in a straight line,/thought that light might, therefore, consist of particles./
3. However, in 1740,/a scientist said if light consists of particles/why do we not feel the particles bouncing on our eyes;/although this question was ignored,/ Young/was nevertheless urged to do his experiment./

The T list

6. Before this,/in 1350,/the mentally disturbed were brutally treated/because it was thought that mental disease was due to being perverse/and possessed of the devil./
2. Before the experiment of/Young,/in 1704 precisely,/Newton/having shown that light travels in a straight line/thought that light might, therefore, consist of particles./
7. This idea was taken so seriously that/in 1710/a book appeared entitled/ the Witch Hammer,/which associated mental disorders with witchcraft./

Results

Recall performance (mean percent correct) by the CT and TC groups on the first task, that is, on the C and T lists respectively, was 34.45 and 23.75 (as is graphically illustrated in Figure 3). Analysis of variance showed that recall of the C list was significantly better ($p < .01$) than recall of the T list.[1]

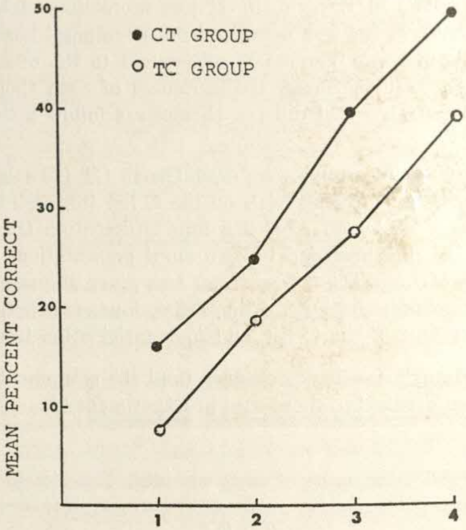


Fig. 3. Mean percent correct recall of segments as a function of trials for groups CT and TC on first lists, Experiment II

Transfer from the thematically organized sentences turned out to be 80.94%, and transfer from the temporally organized sentences was 30.10%. (The transfer measures were calculated using the $E - C/C \times 100$ formula). However, since the CT and TC groups initially differed on degree of learning of the first list (C and T lists respectively), the transfer measures are confounded: it is not known whether original degree of learning of the C and T lists, or the C versus T organization as such, was responsible for the greater transfer shown by the CT group when it was switched from the C to the T list. Recall performance by the CT and TC groups on the second list reached 53.50% and 51.38% respectively, and did not differ significantly ($p > .10$).

The findings on sentence segments are shown in Table 4. This data was analyzed in the same fashion as the sentence-segment data for Experiment I. The results for Experiment are in essential agreement with the results for Experiment I except that the combined mean for proposition recall was significantly higher than the combined mean for name recall ($p < .05$). Date recall was not significantly poorer for group C than group T ($p > .05$), nor the name recall significantly better for group C than group T ($p > .20$). Proposition recall, on the other hand, was significantly better for group C than group T ($p <$ 01).

Table 4. Mean percent correct recall of different sentence segments, Experiment II

	Organization condition		
	C	T	Combined means
Date recall	4.50	11.14	7.82
Name recall	7.90	4.50	6.20
Proposition recall	22.05	8.10	15.75

Figure 4 illustrates recall, by the TC group, of the different sentence segments on both the first and second lists. This figure reflects more dramatically what is shown in Table 4: that when the *same* group is learning a previously learned list in a different organization, recall performance seems to depend heavily on organization. Note how date and proposition recall are affected by the 'T-C switch.' In addition, Figure 4 shows interactions not apparent in Table 4.

DISCUSSION

The superiority of some organizations over others is discussed first, the differential retrieval of various sentence segments subsequently.

Turning to Experiment I, the surprising effect obtained was the lack of significant superiority (in terms of learning performance) of the thematic organization over the name-based organization. This effect is surprising because in the latter organization the sentences of each theme alternate in their order of occurrence in such a way that there is neither a logical nor a temporal organization.

Two reasons can be put forth that might account for this finding. The first is that unlike the sentences in the C list of Experiment II, the sentences in the C list of Experiment I did not resemble usual thematic organizations and did not have embedded in them such phrases as

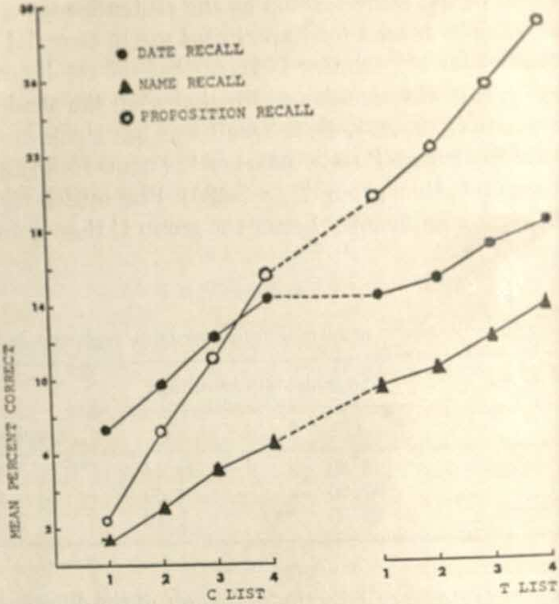


Fig. 4. Mean percent correct recall of the three different types of sentence segments as a function of trials for the TC group on the T and C lists, Experiment II

"because of [a previously described event or name]" and "due to . . . ," phrases which usually signal links between events. Thus, the thematic organization between sentences in Experiment I's C list might have been somewhat 'obscured,' at least as compared to that between sentences in Experiment II's C list.

Secondly, it should be conceded that in the actual, everyday environment, events revolve to a large extent around people and that what people do is 'thematically related'—at least, most of the time. Thus if a subject has to see how some events are generated in the everyday environment, he must have a tendency to organize events around people. And since the actions of individuals are interrelated, he must also, in order to understand what is going on in his world, perceive such relationships or their nature. This means that both a thematic and a 'person' (N) structure of the same material could be established in memory at the same time and that either one could be readily 'translated' into the other. Hence the finding that the thematic and personal organizations are equally strong.

The greater effect, on learning, of the CTW organization in Experiment I argues for the primacy of a thematic memory structure in which there is a hierarchy of events according to time. The findings of Experiment II further suggest that the thematic organization is prepotent over the temporal organization.

Now, inferior recall in one organization as compared to another could be due, first, to less presented material being stored, with retrieval or accessibility potential about the same for the two organizations. (It is assumed here that storage can take place according to only one pattern and not the other.) Or, second, it could be due to less presented material being stored, with retrieval potential in one organization inferior to retrieval potential in the other. Or, third, it could be due to less presented material being accessible under one organization, with amount of items stored about the same in the two organization conditions. (Notice that this last possibility does not mean that storage can take place according to both of two different patterns; what it does mean is that about the same amount of items can be stored, or that the storage pattern is the same, regardless of the organization in which the items are presented. It is certainly not functional to assume that two different storage patterns are established only for one of them to make accessibility inferior.)

Which of the three above-mentioned possibilities is the correct one is not clear from this study. Since there is evidence that about 600 sentences can be stored (Shepard, 1967), perhaps the third possibility is the correct one. This would mean that recall performance in the temporal condition is inferior because retrieval is difficult under that condition, and not because storage is difficult. It should be observed, though, that the suggestion that the different organizations help retrieval rather than storage could be empirically tested by the use of recognition tests. But what, in turn, would underlie the greater difficulty of retrieval in the temporal organization?

The greater retrieval difficulty in the temporal organization (and in the C organization compared to the CTW organization) can be explained if one accepts the theory that there are higher behavioral functions operating in retrieval processes. The thematic organization makes logical and pragmatic relationships between events more obvious than does the temporal organization, and thus permits more 'understanding' of the sentences. In other words, in the thematic organizations (both C and CTW), events are presented in a sequence that is quasi-causal and 'natural'; it fits in with our expectations concerning the flow of events or with our preestablished conceptual habits.

Just *how* such logical and conceptual habits exert control over retrieval processes is unknown. It is possible, however, that by the action of an 'editor' mechanism these higher conceptual habits act as cues during retrieval.

As can be seen in Tables 2 and 4, date recall was affected favorably by organizational schemes in which the dates were temporally aligned, and proposition recall was on the whole (considering the combined means in Table 4) superior under other forms of organization. The first effect, that of date recall, could have any of several explanations. First, the date sequence might have been used as a retrieval plan; that is, a scheme or a network like that shown on the right-hand side of Figure 1 might have been employed—each date coupled with an event, the date acting as a cue and as a link between different pieces of information. Another possibility is that the temporal alignment of dates might have caused the subjects to pay 'more attention' to the dates and that this greater degree of attention was responsible for the superior date performance. And, of course, a combination of the two processes might have been at work. Paying more attention might have resulted in the use of a temporal retrieval plan, or vice versa. This third hypothesis receives some credence from the higher rate of date learning in the CTW organization as compared to the other organizations in Experiment I (see Table 2). Furthermore, the fact that date recall increased and then dropped in Experiment II (see Figure 4) also lends plausibility to this hypothesis. When the subjects were 'switched' from T to C, the recall of propositions and names, but not of dates, increased, which suggests that more attention was being paid to propositions and names. When switching was from C to T, recall of propositions underwent a sudden drop, suggesting that less attention was being paid to propositions. It should be stressed again that no suggestion is being made to the effect that attention is the only factor. Attention might be one of the factors.

That propositions were better recalled than other sentence segments could also have any of several explanations. The storage units for propositions might be different from the storage units for dates and names. If dates and names are stored in more than one unit (e.g., seven-teen-ten, Fres-nel), and if propositions are stored as one unit (see Bever, Fodor, and Garrett, 1966), then higher proposition recall might be due to the subject's having fewer units to remember in order to recall a proposition completely correctly. Stated differently, intraunit interference might account for the slower learning of dates and names, whereas there would be no intraunit interference for propositions.

Again, attention might also be a factor. If the subjects think propositions are 'more important' than dates, they would pay more attention to the propositions. This last suggestion points out the need to determine if subjects do pay more attention to certain parts of sentences than to others. It should also be clear that we need to know more about the memory or storage units of sentences if we are fully to explain the kind of data on retention of sentence segments that was obtained in this investigation.

Notes

This report is based on a thesis submitted in partial fulfillment of the requirements of the degree of master of science at the University of Alberta. The author is deeply indebted to Dr. K. V. Wilson for his indispensable assistance and guidance of this research. He also thanks Drs. W. Rozeboom, W. Runquist, and J. Sampson for their enlightening criticism of the study. His present address is at the Laboratoire de Psychologie expérimental de la Sorbonne, 28 rue Serpente, 75 Paris 6, France. Received for publication March 2, 1970.

1. Further corroborative evidence of this effect was obtained by running two other groups. Groups IC and IT differed from groups TC and CT in that their first task consisted of learning (using the exact same procedure as was used with the latter groups) an irrelevant (I) list on Mediterranean culture. After the end of the first task, either C or T learning took place. Group IC recalled the C list better than group IT recalled the T list ($p < .01$).

References

Bever, T., Fodor, J., and Garrett, M. 1966. Speech perception: Some experimental results for sentences. Paper presented at the International Congress of Psychology, Moscow.

Bousfield, W. A. 1953. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology* 49:229–240.

Kramer, C. Y. 1956. Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics* 12:307–310.

Miller, G. A. 1962. Some psychological studies of grammar. *American Psychologist* 17:748–762.

Shepard, R. N. 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6:156–163.

CUING OF RECALL ORDER IN IMMEDIATE MEMORY WITH DIFFERENT RATES OF PRESENTATION

Larry E. Wood and James V. Hinrichs
University of Iowa

The subjects recalled 10-number lists in the same order as presented or in ascending natural serial order, at a .5-sec or 2-sec rate. They were told the order of recall before (pre-cue) or after (post-cue) list presentation. Presentation rate had little effect on the relationship between recall order and cuing. Serial-order recall was superior to same-order recall and the 2-sec rate was superior to the .5-sec rate, but cuing interacted with both order and rate, producing greater differences for pre-cuing than post-cuing.

In any immediate-memory task that tests retention of multiple-item lists, one of the factors that may affect performance is recall order. If the subjects are required to reproduce the presentation order in addition to recalling the items, the requirements of such a task may be separated, at least logically, into retention of 'item information' and retention of 'order information.' Crossman (1960) has postulated that item and order information can be stored and retrieved separately and that requiring an increase in retention of one type of information results in a reduction of the other. Some support for this hypothesis has been obtained by Crowder (1969), who found significantly better recall under free-recall instructions than when recall was required to be in the same order as presentation.

An attempt to investigate differential retention of item and order information has also been made by Buschke in a series of studies (Buschke, 1967, 1968; Buschke and Hinrichs, 1968) using a procedure that requires the subjects to recall numbers in ascending natural *serial* order. Thus, a list of numbers presented as 18, 12, 1, 20, and 6 would be correctly recalled as 1, 6, 12, 18, and 20. Such a natural serial ordering of numbers is highly overlearned, and the relative order of recall for two particular numbers would be the same on any trial in which they both appear. Therefore, serial-order recall should greatly reduce the necessity for the subject to retain order information in comparison to a condition where recall is required to be in the *same* order as that in which the items were presented.

One of the questions with which Buschke has been concerned is whether differential retention of item information in serial-order and same-order recall should be attributed to differences in encoding at the time of presentation or to differences in retrieval processes. From one study (Buschke, 1967), he concluded that performance differences are largely dependent on differences in storage processes. Furthermore, when subjects were asked for serial-order and same-order recall on two successive presentations of the same list of items, information necessary for same-order recall was stored during presentation on both trials. In a further test of these conclusions (Buschke, 1968), the subjects were asked for serial- or same-order recall, but they were told the type of recall before presentation (pre-cue) on only half the trials. On the remaining trials, they were not informed about the type of recall until after the items had been presented (post-cue). Buschke reasoned that if performance differences between serial- and same-order recall depend on retrieval processes, then cuing of recall order should not differentially affect the two types of recall. The results showed that recall in the pre-cue condition was significantly better for serial ordering than for 'same' ordering; but in post-cuing, no significant difference between the two orderings was found. Buschke concluded that in the pre-cue condition the subject is able to store more item information for serial-order than for same-order recall because he is not required to retain information about order. However, in the post-cue condition he is forced to remember order information on all trials if he is to prepare for possible same-order recall, thereby eliminating the serial-order advantage.

One mechanism that would allow the subject to store more item information for serial-order recall than for same-order recall is the use of differential rehearsal strategies at the time of presentation. To enhance storage of order information for same-order recall, items could be rehearsed in order as they were presented. For serial-order recall, however, each item could be rehearsed by itself as it was presented, thereby maximizing its memory strength (Wickelgren and Norman, 1966) or increasing the likelihood of transfer to long-term storage (Atkinson and Shiffrin, 1968). Then during recall, as suggested by Buschke and Hinrichs (1968), the subject could scan the possible stimulus set in ascending natural serial order, using partial information or variable-decision criteria to determine whether or not a particular item had been presented.

The present study was an attempt to investigate the role of rehearsal strategies in the retention of item and order information by manipulating both recall order and rate of presentation. According to the assumptions outlined above, a slow rate should be optimal for different methods of rehearsal, maximizing differences between 'same' and serial ordering with pre-cuing and minimizing post-curing differences. On the other hand, a fast rate of presentation should tend to limit differential rehearsal strategies and thereby attenuate the effects of pre- and post-cuing of serial-order and same-order recall.

METHOD

—Subjects—The subjects were 40 students from the introductory psychology courses at the University of Iowa and were tested in groups of five.

—Design—The basic design was a 2 × 2 × 2 factorial with one between-subject factor and two within-subject factors. The between-subject factor was presentation rate (one item per .5 sec, or one item per 2 sec), and the within-subject factors were recall order (serial, NS; or same, SO) and cuing of recall order (pre or post). An additional within-subject variable was introduced to allow examination of presentation-position effects in each condition. The 10-item presentation sequence was divided into two halves, with positions 1–5 designated the primacy component and positions 6–10 the recency component.

—Materials—The stimulus materials consisted of 40 lists of 10 of the numbers 1–20 inclusive, constructed from a table of random permutations. Each of the four recall-order and cue-treatment combinations was randomly assigned to 10 of the 40 stimulus lists with the restriction that a particular treatment combination be assigned to no more than 3 successive lists. Three additional sets of lists were constructed by counterbalancing the assignment of conditions to lists so that the sequence of the lists remained the same in each set but each list appeared once under each treatment combination across the four sets. A different subgroup of five subjects within each rate condition received one of the four sets of lists.

—Procedure—The stimulus lists were presented via a tape recorder. In the pre-cue condition, the words "Serial" or "Same" were spoken 2 sec before list presentation to designate the recall order for that trial; the word "Recall" was then spoken 1 sec after the last item. In the post-cue condition, a "Ready" preceded the list by 2 sec; and "Same" or "Serial" was spoken 1 sec after the last item to designate recall order.

The subjects were allowed 20 sec to recall the 10 numbers and to write them in rows of 10 boxes on prepared answer sheets. They were instructed to put the numbers in their correct relative positions from left to right, either in the same order as presented, for same-order recall, or from lowest to highest, for serial-order recall. In producing their responses, they were not required to write each succeeding item from left to right but were allowed to 'skip around.' They were encouraged to try to remember as many items as possible, but they were not required to produce 10 responses on every trial. They were further informed that the four recall-order and cuing combinations would be equally likely to occur on any trial. Four practice lists, one for each treatment combination, were presented prior to the 40 experimental lists, and a 2-min rest period followed the first 20 experimental lists.

RESULTS

Since the study was concerned primarily with retention of item information, the relative position of each item in the response protocol was disregarded in scoring. Correct items that were repeated were counted only once. Table 1 lists the mean total number of correct responses by first and second halves (primacy and recency) for each of the experimental conditions.

The presentation-position curves for each of the experimental conditions are shown in Figure 1. To facilitate comparison of the curves, adjacent pairs of positions were averaged. As the curves illustrate, there was a clear difference between serial-order recall in the pre- and post-cue conditions (Pre-NS and Post-NS), with the latter very simi-

Table 1. Means and standard errors for number of correct responses

Presentation position	Condition			
	Pre-NS	Pre-SO	Post-NS	Post-SO
		Fast rate		
First half				
M	2.840	2.655	2.585	2.390
SE	.131	.158	.140	.183
Second half				
M	3.605	2.895	2.965	2.840
SE	.085	.124	.094	.137
		Slow rate		
First half				
M	3.545	2.850	2.495	2.405
SE	.145	.090	.104	.152
Second half				
M	3.910	3.370	3.205	3.185
SE	.126	.145	.142	.137

lar to the same-order conditions (Pre-SO and Post-SO), which also were quite similar to each other. Serial-order recall in the pre-cue conditions (Pre-NS) was much higher in the middle of the list for both rates, and in the primacy portion of the function for the slower rate; otherwise, there was little difference in the curves as a function of rate, aside from general level of performance.

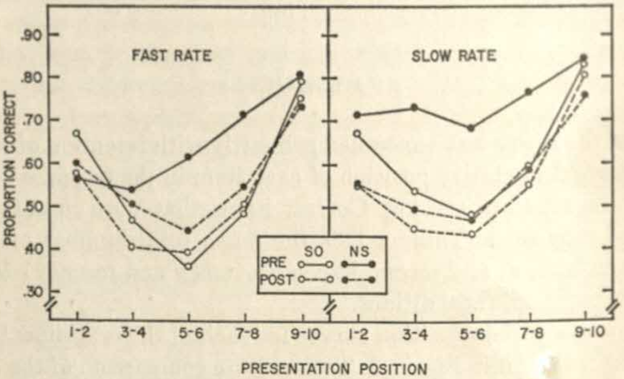


Fig. 1. Proportion of responses correctly recalled as a function of presentation position

The overall similarity of the curves is confirmed by the failure of presentation position to interact with any of the other variables in an overall analysis of variance (all $ps > .05$), although a separate analysis of each rate condition did yield some support for the differences in the Pre-NS curves. For the fast rate, Pre-NS was significantly better than Pre-SO for recency [$F(1, 19) = 28.07$, $p < .001$], but not for primacy [$F(1, 19) = 1.93$, $p > .10$]. At the slow rate, however, Pre-NS was superior to Pre-SO for both primacy [$F(1, 19) = 40.26$, $p < .001$] and recency [$F(1, 19) = 24.30$, $p < .001$]. All the curves show a pronounced recency effect, which resulted in a significant main effect of presentation position [$F(1, 38) = 29.01$, $p < .001$]. The means were 3.08 for recency and 2.62 for primacy.

The failure of the interaction of rate, cue, and recall order to reach an acceptable level of significance [$F(1, 38) = 3.49$, $p > .05$] indicated that the relationship between cuing and recall order was essentially the same for both presentation rates. The nature of this relationship was reflected in a significant interaction of cue and recall order [$F(1,38) = 33.34$, $p < .001$], such that serial-order recall was substantially better than same-order recall in pre-cuing [$F(1,38) = 104.93$, $p < .001$] but only slightly better in post-cuing [$F(1,38) = 4.21$, $p < .05$]. For pre-cuing the means were 3.48 and 2.94 for serial- and same-order recall respectively; for post-cuing they were 2.81 and 2.71. Separate analysis of each rate condition yielded somewhat contrary results: for the slow rate, serial- and same-order recall did not differ with post-cuing ($F < 1$), while for the fast rate, serial-order recall was slightly better than same-order recall with post-cuing [$F(1,19) = 4.47$, $p < .05$]. Further analysis of the interaction of cue and recall order collapsed across rate revealed that pre-cuing was better than post-cuing for both same-order recall [$F(1,38) = 20.70$, $p < .001$] and serial-order recall [$F(1,38) = 162.36$, $p < .001$].

Cuing of recall order also interacted with presentation rate [$F(1,38) = 15.51$, $p < .001$]. The means for the slow and fast rates with pre-cuing were 3.42 and 3.00 respectively, and 2.82 and 2.70 with post-cuing. Individual comparisons between means revealed that pre-cuing was significantly better than post-cuing both at the fast rate [$F(1,38) = 33.29$, $p < .001$] and at the slow rate [$F(1,38) = 128.38$, $p < .001$]. However, the decrement in performance with post-cuing was greater at the slow rate, so that the slow rate's superiority in the pre-cue condition [$F(1,38) = 63.97$, $p < .001$] was significantly reduced in the post-cue condition [$F(1,38) = 5.84$, $p < .025$].

DISCUSSION

The present study was concerned with the role of different rehearsal strategies in storage of item and order information for serial-order and same-order recall. The finding that serial-order recall was markedly superior to same-order recall for pre-cuing but not for post-cuing at both rates is strong evidence that the differences between the two recall orders are largely due to different storage processes rather than to different retrieval processes. For if the superiority of serial-order recall were due to a more effective retrieval process, then that superiority should have been maintained even when the subject was not informed about the type of recall until after list presentation. It could be argued that serial-order recall is a more effective retrieval process but that because the subject uses a rehearsal strategy in post-cuing that is appropriate for same-order recall, the order cues interfere with the serial-order retrieval process, reducing its effectiveness. But even if this argument were valid, serial-order recall would still be expected to have been better than same-order recall in post-cuing at the faster presentation rate, because differential rehearsal would be minimal. While some evidence for greater serial-order recall was obtained in the separate analysis of the fast-rate condition, the difference in post-cuing was very small compared to the pre-cuing difference.

Even though better retention for serial-order recall appears to be largely due to storage processes, the present data offer little support for different rehearsal strategies as a means of accounting for the superiority. At the fast presentation rate, opportunity for rehearsal would be minimal; yet pre-cuing resulted in substantially better serial-order recall than same-order recall, whereas post-cuing did not. If differential rehearsal were the major factor leading to the serial-order superiority, then the two cue conditions should have been much less disparate. Whatever mechanism did operate to produce better storage of item information for serial-order recall, it is clear from the curves in Figure 1 that its major effect was to increase storage of information about items in the middle of the list.

Because the presentation-position curves for the post-cue serial-order conditions were quite similar to the curves for the same-order conditions, it seems reasonable to assume that the subjects tended to use the storage process appropriate for same-order recall in the post-cue condition. However, this conclusion must be tempered by the finding that performance with the slow rate was only slightly better than that with the fast rate in post-cuing but substantially better for pre-cuing, regardless of the recall order. It might be expected that a slower rate should be equally advantageous for same-order recall in both pre- and post-cuing if the subject always prepares for same-order recall in the post-cue condition.

An interesting aspect of the presentation-position curves was the pronounced recency effect in all conditions. This is somewhat surprising because a recall cue ("Same," "Serial," or "Recall") occurred at the end of each list, essentially forming a stimulus suffix. Typically, a stimulus suffix tends to markedly reduce recall performance in the last two or three positions (see Crowder, 1967; Dallet, 1965), which clearly was not the case in the present study.

One final result that warrants comment was the finding that pre-cuing was significantly better than post-cuing across both rate conditions and both recall orders. The decrement in performance for the post-cue condition may be due to the additional information contained in the recall cue. In pre-cuing the "Recall" is merely a signal to begin recall for which the subject has previously prepared; whereas there is uncertainty associated with the "Same" or "Serial" cue in the post-cue condition, an uncertainty which must be resolved. This uncertainty could be related to the evidence that information processing between the presentation and the recall of a list of items leads to reduction in the level of retention (Posner and Rossman, 1965).

Notes

The first author was supported by a National Defense Education Act predoctoral fellowship. Additional support was provided by National Institute of Mental Health Grant MH-16362 to the second author. Received for publication September 7, 1970.

References

Atkinson, R. C., and Shiffrin, R. M. 1968. Human memory: A proposed system and its control processes. In *The psychology of learning and motivation*, ed. K. W. Spence and J. T. Spence, vol. 2, pp. 90–195. New York: Academic Press.

Buschke, H. 1967. Two kinds of short-term storage. *Psychonomic Science* 8:419–420.

Buschke, H. 1968. Perceiving and encoding two kinds of item-information. *Perception and Psychophysics* 3:331–336.

Buschke, H., and Hinrichs, J. V. 1968. Controlled rehearsal and recall order in serial list retention. *Journal of Experimental Psychology* 78:502–509.

Crossman, E. R. F. W. 1960. Information and serial order in human immediate memory. In *Information theory* [Proceedings of the fourth London conference on information theory], ed. C. Cherry, pp. 147–159. London: Butterworth.

Crowder, R. G. 1967. Prefix effects in immediate memory. *Canadian Journal of Psychology* 21:450–461.

Crowder, R. G. 1969. On free recall and free-recall scoring, in immediate memory. *Psychonomic Science* 14:255–256.

Dallett, K. 1965. "Primary memory": The effects of redundancy upon digit repetition. *Psychonomic Science* 3:237–238.

Posner, M. I., and Rossman, E. 1965. Effect of size and location of informational transforms upon short-term retention. *Journal of Experimental Psychology* 70:496–505.

Wickelgren, W. A., and Norman, D. A. 1966. Strength models and serial position in short-term recognition memory. *Journal of Mathematical Psychology* 3:316–347.

TRANSFER AS A FUNCTION OF PARADIGM AND INTERSPERSED-PAIR DIFFICULTY

G. J. Johnson
University of British Columbia, Canada

Transfer performance for AB-AB′, AB-CD, and AB-AC paradigms was investigated with subset items that were less difficult than, more difficult than, or of the same paradigm type as the test pairs. Performance on the three test paradigms was a decreasing function of the difficulty of interspersed pairs, and intralist-intrusion data suggest that the strength of intrusion tendencies generalized to a given set of test pairs increases with the difficulty of the items with which the set is mixed. Difficulties in interpreting the results of studies using the mixed- versus unmixed-list manipulation to test assumptions about the use of systematic strategies in verbal learning are discussed.

Recent studies of transfer effects in paired-associate learning have shown that performance on items representing a given paradigm type may vary as a function of whether the items are presented in a mixed list or in an unmixed list. Several factors have been isolated as possible determinants of such discrepancies. Postman (1966) emphasized the differential opportunity provided subjects under the two conditions to apply a given strategy consistently to all items in a list. That is, performance on transfer pairs such as AB-AB′ (stimuli identical, responses synonymous) may be facilitated with an unmixed list by the availability of an easily detectable and obviously effective list rule. With a mixed list, even if the appropriateness of such a rule is detected, the efficiency with which it may be applied is limited, since the subject must differentiate from other items those to which the particular rule is applicable.

A very similar interpretation was offered by Paul and Paul (1968), in connection with the concept of transfer-activated response set (TARS). In an unmixed list the consistent presentation of AB-AB′ pairs is assumed to activate a generalized response set whose effect is to facilitate performance by restricting the subject's responses to synonyms of first-list response terms. In a mixed list such a set is assumed to be either not elicited or not consistently reinforced due to the inappropriateness of the set with regard to the other, not AB-AB′ pairs.

Such 'cognitive strategy' interpretations of the transfer effects of mixed and unmixed lists have drawn on findings of a mixed-list performance decrement on the type of item (say, AB-AB′) which, when presented in an unmixed list, may be expected to provide an effective rule of response restriction (Postman, 1966; Wickens and Cermack, 1967). However, a number of studies suggest that the importance of the availability of a consistent rule with an unmixed list may have been overestimated relative to other factors, for superior performance on certain items under the mixed-list condition has been frequently observed (Johnson, 1970a; Johnson and Penney, 1965, 1966; Paul and Silverstein, 1968; Slamecka, 1967; Underwood and Schulz, 1960; Wickens and Cermack, 1967).

Johnson (1970a) noted a general pattern in the results of studies of the effects of mixed and unmixed lists—a pattern which suggests that performance on a given set of pairs varies inversely with the difficulty (directly with the initial associative strength) of the items with which the test pairs are presented. Performance on pairs with the 'easier' mixed-list subsets tends to be superior to that with an unmixed list, whereas performance on pairs with the more 'difficult' mixed-list subsets tends to be inferior to that with an unmixed list.

This pattern was interpreted in terms of the assumption that the strength of intralist intrusion tendencies potentially interfering with performance on a given set of test items varies directly with the relative difficulty of the pairs with which they are mixed. An obvious implication of this interpretation is that, for any given paradigm type, the *direction* as well as the magnitude of differences between the mixed and unmixed lists may vary as a function of the specific composition of the mixed list.

Previous studies using the manipulation of mixed- and unmixed-list conditions to assess the role of systematic strategies in transfer performance have confounded the consistency of a given rule's applicability in the two conditions with differences in the difficulty of the items with which the test pairs are presented under the two conditions. Significantly superior performance with an unmixed list has been obtained *only* where the mixed lists involved included subsets of items of greater difficulty than the test paradigm (Johnson, 1970a;

Postman, 1966; Slamecka, 1967; Wickens and Cermack, 1967). But if such results reflect primarily the greater availability of a consistent rule in the unmixed list, then that facilitation should also obtain when the level of difficulty of other subsets in the mixed list is equal to or less than that of a given test paradigm.

The present study investigated transfer performance for three different paradigms, with three levels of difficulty for the pairs with which the test items were presented. Items representing the paradigms AB-AB′, AB-CD, and AB-AC were presented (*a*) in a mixed list that contained items of *less* difficulty than the test pairs, (*b*) in an unmixed list where all items were of a *similar* level of difficulty, and (*c*) in a mixed list that included items of *more* difficulty than the test pairs. The two mixed-list conditions do not appear to differ with respect to the consistency with which the rules appropriate to test pairs are applicable nor with respect to the degree to which a transfer-activated response set elicited by a given set of test items may be activated. To the extent that the rule appropriate to a given paradigm provides the unmixed-list subjects with an effective guide to performance, the unmixed list should yield the highest level of performance, while the mixed-list conditions should not differ. But if transfer performance is directly related to the initial associative strength of the pairs with which the test items are mixed, the 'easy' mixed list should yield the highest, the unmixed list an intermediate, and the 'difficult' mixed list the lowest level of performance.

METHOD

—Materials—Each subject learned two lists of 12 two-syllable adjectives. The transfer list was the same for all subjects, while first-list items were selected so as to provide the various experimental conditions. Fifteen first-list conditions were used. Items representing each of the paradigm types AB-AB′, AB-CD, and AB-AC were presented under an unmixed-list and the two mixed-list conditions. That is, in the mixed-list AB condition six test pairs of each paradigm type were mixed with six pairs carried over from the first list (AB-AB pairs). And in the mixed-list ABj condition the test pairs were presented along with six items representing an AB-AB jumbled (AB-ABj) paradigm. For the latter items stimulus-response roles were reversed for half of the pairs and, further, all associates were re-paired across lists.

In order to have all items in the unmixed-list condition presented under the mixed-list conditions, two different lists were constructed for each mixed list. For example, in one AB-AB′/AB-AB, subset A of the transfer task represented AB-AB′ items, while subset B appeared as AB-AB pairs. In the second AB-AB′/AB-AB, the subset-paradigm relation was reversed. This procedure was used in constructing first lists for each of the two mixed-list conditions with each of the three test paradigms.

The adjectives used in constructing the lists were taken from Haagen (1949) and from Melton (1940). With the exception of the synonymous B-B′ relation for AB-AB′ pairs and that of identity for AB-AB pairs, all items were selected to minimize meaningful and formal similarity both within and across first and second lists. The lists were presented on an MTA Scholar at a 2:2 rate with a 4-sec intertrial interval. Three different random orders of the 12 items in each list were used to reduce serial effects.

—Subjects—The subjects were 270 volunteers from eleventh- and twelfth-grade classes in the Vancouver public schools. None had previously participated in a verbal-learning experiment. With one unmixed-list condition and four mixed-list conditions for each of the three paradigms, there were 15 groups of 18 subjects each. Assignment to conditions was carried out in blocks of 15 with one subject per block being assigned to a given condition. Within each of the blocks assignment followed a different predetermined random ordering of the 15 conditions. All subjects served individually.

—Procedure—Each subject was given a maximum of 25 anticipation trials on the first list. First-list training was terminated when he achieved a criterion of one perfect recitation of the list or after 26 presentations of the list if he did not achieve criterion. Following first-list learning, he was given 10 trials (excluding the first exposure) on the transfer task. The intertask interval was approximately 2 min.

RESULTS

Comparison of the 15 groups in terms of trials to criterion on the first list was used as a means of assessing possible differences in initial ability. Mean number of trials to criterion for all subjects was 15.18, while individual means for the 15 groups ranged from 12.16 to 18.27 trials. Analysis of variance indicated no significant differences among conditions [$F(14, 255) = 1.22$].

Performance on the transfer task was examined in terms of number of correct anticipations over 10 trials. Table 1 summarizes these data for each of the three test paradigms (AB-AB′, AB-CD, and AB-AC) under each of the three list conditions (mixed-list AB, unmixed list, and mixed-list ABj). The measures under the unmixed-list conditions represent the performance of 18 subjects on 12 items for each test paradigm. The comparable measures for the mixed lists represent the performance of 36 subjects on one of two subsets of 6 items each. For each test paradigm, data from the two subsets within each mixed-list condition were combined by ranking the subjects within each of the two groups for a given condition on the basis of their trials to first-list criterion. For equally ranked subjects, data from the two subsets of transfer items were combined and treated as having resulted from the performance of individual subjects. This arrangement resulted in nine groups of 18 subjects each, making possible a factorial analysis of paradigm and list effects.

Table 1. Mean correct responses for AB′, CD, and AC pairs under three list conditions

	Mixed-list AB	Unmixed list	Mixed-list ABj
AB′	89.83	86.55	79.00
CD	90.11	85.50	79.72
AC	75.39	63.05	55.89

The data in Table 1 show little or no difference between performance on AB-AB′ and AB-CD pairs. Mean numbers correct for AB-AB′ and AB-CD pairs respectively were 85.13 and 85.11. This lack of positive transfer for AB-AB′ items relative to control pairs is in line with Postman's results (1966). Substantial negative transfer effects were apparent in performance on AB-AC pairs, which yielded 64.78 mean correct responses. By Dunnett's test, the amount of negative transfer was reliable at the .01 level. With regard to the list variable, it may be noted that mean correct anticipations for each of the test paradigms were highest for the mixed-list AB condition, intermediate for the unmixed-list condition, and lowest for the mixed-list ABj condition. This decrease in performance as a function of increased level of difficulty of the pairs with which the test items were mixed was highly significant [$F(2, 153) = 6.49$; $p < .001$]. Differences in performance attributable to the list variable did not vary as a function of paradigm type [$F(4, 153) < 1$].

Table 2 presents mean numbers of correct responses on AB-AB and AB-ABj items in mixed lists containing pairs representing each of the three test paradigms. Each entry was obtained from the performance of 36 subjects on one of the two subsets for each paradigm. When performance is compared across the mixed lists in which these pairs

Table 2. Mean correct responses for AB and ABj pairs under three mixed-list conditions

	Mixed-list AB′	Mixed-list CD	Mixed-list AC
AB	102.33	109.28	100.72
ABj	51.94	45.72	42.33

appeared, a pattern similar to that revealed by performance on the three test paradigms as a function of difficulty of interspersed pairs is apparent. When mixed with AB-AB′, AB-CD, and AB-AC pairs respectively, mean numbers correct for AB-AB and AB-ABj items combined were 81.92, 78.02, and 71.53. An exception to this pattern is the superior performance on AB-AB items in the mixed-list CD condition. The overall F for differences among the three mixed-list conditions (AB′, CD, and AC) was not significant [$F(2{,}210) = 1.98$]. However, since AB-AB′ and AB-CD items did not differ in difficulty, mixed-list AC was compared to mixed-list AB′ and CD combined. This comparison showed that the decrement in performance on AB-AB and AB-ABj pairs under the mixed-list AC condition was significant [$F(1, 210) = 3.95$; $p < .05$].

Interlist intrusion errors

An analysis of errors on the transfer task attributable to correct associations for first-list learning was carried out for AB-AB′ and AB-AC pairs as a function of the lists in which they were presented. For AB-AB′ pairs 6.78 mean first-list intrusions occurred under the mixed-list AB condition, 1.22 under the unmixed-list condition, and 1.89 under the mixed-list ABj condition. The AB-AC items yielded corresponding values of 6.00, .44, and .55. Analysis of variance indicated that differences among the three list conditions were reliable [$F(2, 102 = 35.15$; $p < .001$]. The Newman-Keuls test showed that the separation between the mixed-list AB condition and each of the other two list conditions was significant beyond the .01 level. No further significant effects were indicated.

For AB-ABj items, the number of transfer-list errors corresponding to correct responses in first-list learning was compared across mixed-list AB′, CD, and AC conditions. The greatest number of such errors occurred under CD, although the difference among the three conditions did not achieve statistical significance [$F(2, 51) = 2.08$; $p > .05$]. Mean first-list intrusions on transfer items were 4.44 for CD, 3.44 for AB′, and 2.50 for AC.

Intralist intrusion errors

Mean numbers of errors attributable to interpair generalizations during second-list learning under the unmixed-list conditions were 6.22 for AB-AB′, 4.61 for AB-CD, and 9.61 for AB-AC pairs

[$F(2, 51) = 4.15$; $p < .025$]. The major source of variance among the three paradigms was due to the separation between AB-AC pairs versus AB-AB′ and AB-CD pairs combined [$F(1, 51) = 7.48$; $p < .01$]. The difference between AB-AB′ and AB-CD pairs was not reliable ($F < 1$).

Table 3 presents number of intralist generalization errors for each of the six mixed-list conditions. For each mixed list, intrusion errors are classified according to their origin (generalizations from) and their place of occurrence (generalizations to). Each entry is based upon the performance of 36 subjects. The AB-ABj entries exclude intralist intrusions that might be attributable to associations acquired during first-list learning and thus perhaps underestimate the strength of generalization tendencies within that subset of pairs. Data from each of the mixed-list conditions were analyzed separately by comparing the two paradigm types in terms of number of 'generalizations to' and number of 'generalizations from.' Each of the *F*s involved is based upon 1 and 35 *df*. The major purpose of these analyses was to determine whether there is a systematic relationship between the relative difficulty of two subsets of pairs presented in a mixed list and the distribution of intralist generalization errors. The general pattern that emerges is fairly clear. The more difficult of two subsets of mixed-list pairs tended to be the more frequent source of, as well as the more frequent place of occurrence of, intralist generalizations.

For number of 'generalizations to,' this relationship was significant for mixed lists AB-AB′/AB-ABj ($F = 22.24$; $p < .001$), AB-CD/

Table 3. Number of generalization errors within each mixed list

	AB-AB′/AB-AB				*AB-AB′/AB-ABj*		
	To AB′	To AB	Total		To AB′	To ABj	Total
From AB′	43	51	94	From AB′	27	56	83
From AB	80	34	114	From ABj	50	95	145
Total	123	85	208	Total	77	151	228
	AB-CD/AB-AB				*AB-CD/AB-ABj*		
	To CD	To AB	Total		To CD	To ABj	Total
From CD	100	6	106	From CD	77	76	153
From AB	3	39	42	From ABj	38	152	190
Total	103	45	148	Total	115	228	343
	AB-AC/AB-AB				*AB-AC/AB-ABj*		
	To AC	To AB	Total		To AC	To ABj	Total
From AC	92	57	149	From AC	48	94	142
From AB	58	22	80	From ABj	92	116	208
Total	150	79	229	Total	140	210	350

AB-AB ($F = 15.07$; $p < .001$), AB-CD/AB-ABj ($F = 16.36$; $p < .001$), AB-AC/AB-AB ($F = 9.24$; $p < .005$), and AB-AC/AB-ABj ($F = 7.46$; $p < .01$). For AB-AB′/AB-AB $F = 2.52$; $.25 > p > .10$. In terms of 'generalizations from,' a significantly greater number of errors was associated with the more difficult subset in mixed lists AB-AB′/AB-ABj ($F = 8.11$; $p < .01$), AB-CD/AB-AB ($F = 13.61$; $p < .001$), AB-AC/AB-AB ($F = 7.84$; $p < .01$), and AB-AC/AB-ABj ($F = 12.05$; $p < .001$). In AB-CD/AB-ABj the relationship was not significant, although the results were in the direction of the general pattern. The single reversal of this pattern occurred in AB-AB′/AB-AB, where a greater number of intrusions was generated by AB pairs than by AB′ pairs ($F < 1$).

For the six mixed lists combined, those items classifiable as the more difficult of two subsets yielded a highly significantly greater number of 'generalization to' errors [$F(1, 210) = 59.62$; $p < .001$] as well as 'generalization from' errors [$F(1, 210) = 22.20$; $p < .001$]. Finally, the numbers of intrusion errors of each type associated with the five different paradigms in the six mixed lists closely paralleled the relative level of difficulty of the pairs. Respective mean values for AB, AB′, CD, AC, and ABj items for 'generalizations to' were .96, 1.39, 1.51, 2.01, and 2.73. For 'generalizations from,' the corresponding values were 1.08, 1.22, 1.79, 2.02, and 2.52.

DISCUSSION

Performance on the three major test paradigms—AB-AB′, AB-CD, and AB-AC—was inversely related to the level of difficulty of the pairs with which the test items were presented. This result implies that attempts to arrive at a general statement of the differences between mixed-list and unmixed-list transfer effects for any given paradigm are likely to be unfruitful. The direction of performance differences with the two kinds of lists was shown to be dependent on the specific composition of the mixed list. That is, performance on test pairs in a mixed list containing a subset of *less* difficult items was better than performance with an unmixed list. Conversely, performance on test items presented in a mixed list containing pairs *more* difficult than the test items was inhibited relative to performance with an unmixed list. Further evidence that performance on a given subset of pairs in a mixed list varies directly with the initial associative strength (inversely with the difficulty) of interspersed items was provided by the results on AB-AB and AB-ABj pairs. Performance on these pairs when they were mixed with AB-AC items was inhibited relative to that when they were mixed with AB-AB′ or AB-CD pairs.

Interlist-intrusion data revealed that the presence of AB-AB items in a mixed list significantly increased the frequency or errors, on AB-AB′ and AB-AC pairs, attributable to first-list associations. This finding parallels evidence cited by Paul and Paul (1968) in support of the facilitative effects of unmixed-list presentation on transfer performance. That is, in unmixed lists AB-AB′ or AB-AC a transfer-activated response set apparently functions so as to inhibit first-list associations. In mixed-list AB, the response set serving to inhibit first-list responses apparently is either not activated or conflicts with one that enhances the emission of first-list responses.

Postman's (1966) analysis emphasizes a similar factor assumed to operate in favor of performance on the unmixed list. Where one of the mixed-list subsets consists of items (such as AB-AB pairs) for which there is a readily detectable and obviously effective rule, the pattern of intrusion errors is expected to reflect inappropriate generalization of this rule to other mixed-list subsets. The interlist-intrusion data of the present study are closely in line with this expectation. It is clear that the presence of AB-AB pairs in a mixed list decreases the distinctiveness of first- versus second-list associations. However, it is equally clear that this reduction of interlist discriminability did not result in inferior performance on those items mixed with AB-AB pairs. That is, in terms of overall performance, mixed-list AB was superior to the unmixed list for each test paradigm.

The intralist-intrusion data appear to be consistent with the assumption (Johnson, 1970a) that the degree of interitem interference generated by a given set of pairs is a direct function of level of difficulty of the individual items. Comparison of the number of intralist intrusions across the unmixed-list conditions showed a direct relation between the frequency of occurrence of such errors and the relative level of difficulty of the paradigms involved. It should be noted, however, that this pattern of differences parallels that indicated by *total* errors in each of the unmixed lists and thus may not provide a satisfactory assessment of the assumed effects of intralist interference as distinct from those of list difficulty per se. Percentage of errors attributable to intralist intrusions appears to be an equally unsatisfactory measure. Comparison on this basis involves the questionable assumption that frequency of errors of other types (omissions, for example) is unrelated to differences in level of intralist interference. A similar problem of interpretation is presented by the finding that a significantly greater number of within-list intrusions occurred on items of the more difficult of two mixed-list subjects. This result may reflect a tendency for performance on those pairs to be more susceptible to the effects of interitem interference than is performance on items of the less difficult subset. Assessment of such a contention, however, requires a design that effectively separates the effects of intralist sources of interference from those of inter- or extralist sources.

Data regarding the *source* of generalization errors in the mixed lists more clearly support the hypothesis that the strength of intrusion tendencies generated by a given set of items is a function of level of difficulty of the pairs. Regardless of the place of occurrence of intralist intrusions (and the level of difficulty of pairs on which they occurred), a significantly greater number of intrusions was generated by the more difficult of two mixed-list subsets. Further, when intrusions were classified according to individual paradigms across all mixed-list conditions, the frequencies of intrusions originating from the respective pairs followed very closely the order of difficulty of the five paradigm types. A similar trend may be noted in Postman's (1966) intrusion data for four paradigms presented under a single mixed-list condition.

The one exception in the present results to this pattern of generalization errors in mixed lists may be worthy of note. In AB-AB′/AB-AB, a greater number of intrusions was generated by AB than by AB′ pairs. Differential response availability for AB and AB′ seems to provide an obvious explanation. However, such an interpretation is inconsistent with the intrusion data of AB-AC/AB-AB, where differences in response availability between the two subsets are even greater. In the latter case significantly more intrusions were generated by AC than by AB pairs.

The results associated with interpair intrusions in AB-CD/AB-AB reflect the operation of a factor that may generally be expected to facilitate performance. It has been suggested elsewhere (Johnson, 1970b) that a high degree of intersubset discriminability may limit intralist intrusion tendencies to those generated among items within each mixed-list subset. The distinctiveness between items carried over from the first list and those presented only in the transfer task appears to have been relatively complete, since the intrusion data indicated a virtual absence of generalizations across AB and CD subsets. The somewhat superior performance on AB pairs mixed with CD items relative to that on AB pairs mixed with AB′ items follows from these considerations. The importance of the high associative strength of AB pairs in determining intersubset differentiation is illustrated by a comparison of generalizations across AB and CD versus those across ABj and CD subsets. In terms of a first- versus second-list distinction, intersubset discriminability for the two mixed lists is constant. Yet the low associative strength of ABj pairs appears to have resulted in a relatively substantial degree of intersubset generalization.

In sum, the present results are in general agreement with the conclusions reached by Paul and Paul (1968). Transfer effects for a given paradigm were clearly dependent on characteristics of the items with which the test pairs were presented in mixed versus unmixed lists. The present data further indicate that the level of difficulty of interspersed pairs is a major factor determining the degree and direction of differences between the two kinds of lists for a given test paradigm. This finding does not necessarily conflict with previous assumptions about the influence of systematic strategies on transfer performance. But it does focus attention on the difficulties involved in interpreting the results of studies that use the manipulation of mixed- and unmixed-list conditions to test such assumptions. For if the designs employed fail to control for possible effects of differential variation in intralist difficulty, inferences about the utilization of cognitive strategies under various conditions of transfer are apt to be misleading.

Notes

Received for publication September 18, 1970.

References

Haagen, C. H. 1949. Synonymity, vividness, familarity, and association value rating of 400 pairs of common adjectives. *Journal of Psychology* 27: 453–463.

Johnson, G. J. 1970a. Systematic strategies and transfer effects with mixed- and unmixed-list designs. *Psychological Reports* 26:475–484.

Johnson, G. J. 1970b. Effects of item discriminability in mixed-list presentation of paired associates. *Psychonomic Science* 21:371–372.

Johnson, G. J., and Penney, R. K. 1965. Transfer effects of mixed and unmixed list designs in paired associate learning of children. *Psychonomic Science* 2:171–172.

Johnson, G. J., and Penney, R. K. 1966. The effect of mixed vs. unmixed list designs in paired associate learning. *Journal of Verbal Learning and Verbal Behavior* 5:234–239.

Melton, A. W. 1940. Materials for use in experimental studies of learning and retention of verbal habits. Mimeographed manuscript, University of Missouri.

Paul, C., and Paul, H. 1968. Transfer activated response sets in verbal learning and transfer. *Psychological Review* 75:537–549.

Paul, C., and Silverstein, A. 1968. Relation of experimentally produced interlist intrusions to unlearning and retroactive inhibition. *Journal of Experimental Psychology* 76:480–485.

Postman, L. 1966. Differences between unmixed and mixed transfer designs as a function of paradigm. *Journal of Verbal Learning and Verbal Behavior* 5:240–248.

Slamecka, N. J. 1967. Transfer with mixed and unmixed lists as a function of semantic relations. *Journal of Experimental Psychology* 73:405–410.

Underwood, B. J., and Schulz, R. W. 1960. Response dominance and rate of learning paired associates. *Journal of General Psychology* 62:153–158.

Wickens, D., and Cermack, T. S. 1967. Transfer effects of synonyms and antonyms in mixed and unmixed lists. *Journal of Verbal Learning and Verbal Behavior* 6:832–839.

Corrections

The following alterations are to Dr. George Sperling's article on "Binocular Vision: A Physical and a Neural Theory" (Vol. 83, No. 4). On page 479, Equation 8, line 1, ∇^n should be ∇^2. On page 479, Equation 9, line 1, ∇^2 should be ∇^n. On page 484, paragraph 1, the last line, $\Delta v < 0$ should read $\Delta u < 0$. On page 527, Equation A-1 and the equation immediately below it should both read $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}$, instead of with the minus signs omitted. On page 531, Equations B-11 and B-12 should read

$$y_i = \frac{x_i + \alpha y_i}{1 + k y_j} \quad \text{and} \quad y_i = \frac{(\beta x_i)}{1 + (\beta k) y_j} = \frac{x'}{1 + k' y_j}$$

respectively, instead of with incorrect subscripts. On page 532, line 7, (k) should read (k').

BOOK REVIEWS

Julian Hochberg
Columbia University

Principles of Perceptual Learning and Development

By Eleanor J. Gibson. New York: Appleton-Century-Crofts, 1969. Pp. viii, 537. $8.50.

Winner of the annual Century Psychology Series Award, this book marks two important firsts. It is the first full-length work to be devoted to the marriage of perception and learning theory; and it is the first book by Eleanor J. Gibson, whose research studies in these areas and in the experimental investigation of early child development have established her reputation as a distinguished authority in this newly emerging field in psychology.

Although William James recognized the importance of perceptual learning in motor accomplishments such as billiards and rifle shooting and in professional testing and buying—wine tasting, for example—the factors on which these skills depend have been little investigated. How does the ornithologist come to discriminate and identify many different bird songs, or the psychologist at the animal-behavior station learn to recognize his goat in the herd? How does the surgeon learn to read X-ray plates or, indeed, to improve the fine sensory discrimination required for delicate surgery?

By focusing attention on these questions, Eleanor Gibson has exposed new frontiers of our ignorance and pointed the way to productive new research. The existing evidence is scattered and unsystematic. In this book Professor Gibson has accomplished the feat of bringing together a large number of studies on perceptual learning and development from a variety of sources and of bringing them to bear on a central theoretical viewpoint. Her consistently argued thesis is that perceptual learning consists in "the increased ability of an organism to get information from its environment as a result of practice with the array of stimulation provided by the environment" (p. 77). The "discovery of distinctive features and structure in the world" describes the process—one that is self-regulating in that it achieves the goal of "extracting and reducing the information in stimulation without the necessity of external reinforcement" (p. 4).

The book is neatly divided into three parts. The first part deals with theories of perceptual learning; the second part reviews illustrative experiments, evaluating methods and raising questions for research; and the third part covers perceptual development, both phylogenetic, including imprinting, and ontogenetic, from early infancy.

Capsule summaries of traditional and contemporary theories of perceptual learning provide brief critical abstracts of the views of key investigators. Helmholtz, Titchener, James, Koffka, Carr, and Werner represent traditional theorists; while Brunswik, Ames, Bartlett, Piaget, and Bruner are some of the contemporary psychologists against whose positions the author advances her own differentiation theory that "there is information in stimuli to be picked up by a sensitive, exploring organism" (p. 75).

Curiously, the work of Skinner and his followers is omitted, although they have made discrimination a focal point of the experimental analysis of behavior. Gibson ignores Skinner's distinction between discrimination and differentiation, although it seems useful to consider *discrimination* between stimuli as the process whereby a given response is made to one of a pair of stimuli and not to the other, and *differentiation* between responses as the process whereby the form of the response itself or its topography is altered by differential reinforcement. By failing to include any consideration of such operant-conditioning experiments as Terrace's ingenious work on errorless discrimination, or of related, important applications of 'fading' techniques to education, the book is less comprehensive than could be desired.

Professor Gibson raises two questions for learning theory to answer. First, what is learned? Second, what are the principles for learning? The chapters devoted to differentiation theory carefully elaborate the answer to the first question. With untiring concern to drive the point home, Gibson describes what is learned as the detection of properties, patterns, and distinctive features—properties that may be buried in a welter of impinging stimulation and must be extracted or perceptually isolated.

As the result of extracting patterns of distinctive features and invariant relations not previously responded to, generalization and variability decrease, and response becomes increasingly precise. As opposed to an enrichment theory of perceptual development, a differentiation theory emphasizes the correspondence that develops between variables of stimulation and perception. That the changes in perception which result from experience and practice are functionally adaptive seems to be a reasonable expectation; but unfortunately, bad habits can be learned as well as good ones, and familiarity can lead to proofreader's error as well as to improved discrimination. The assertion that learning necessarily leads to a better correlation between behavior and the environment causes difficulty and fails to consider large categories of phenomena: illusions, negative aftereffects of prolonged stimulation, cases of overconstancy, and maladaptive modifications that often enough do result from interaction with the sources of stimulation. If these phenomena are excluded from the fields of perceptual learning and development, then what processes or mechanisms must be invoked to account for them? An even more important difficulty for the correspondence theory arises from the problem of explaining those changes in perception for which we have no known regular stimulus correlates. How can a correspondence theory account for those changes that are characterized as aesthetic or attitudinal: the acquisition of new tastes and the greater appreciation of subtle qualities of strength, humor, or beauty in art, music, literature, and other human beings? What stimulus features or invariant properties constitute the objective physical correlates for these perceptions?

Professor Gibson, of course, is aware of these problems and is at some pains to extend the scope of her theory to cover a broad range from wine tasting to the discrimination of letters and higher-order symbolic structures. She considers five different sources of stimulation: objects, space, events, representations of these, and coded stimuli (speech, writing, Morse code.)

In every case, the book stresses the discovery of new features that comes with increased familiarity with stimulus variations for which the individual had initially few relevant coding operations. The argument is most convincing when it can specify the stimulus attributes available for discovering; it is less successful in accounting for aesthetic learning.

The question of how the organism learns what a distinctive feature is, or how stimulus dimensions come to be perceptually isolated, is as important as it is difficult to answer. Chapter 6 offers an outline of principles and mechanisms of perceptual learning that is especially concise and persuasive. This chapter presents several propositions that follow from the assumption that in a multidiscrimination learning task it is the dimensions of difference between the members of a set that are learned. (How the set initially got itself established in perception is not considered.) Learning is facilitated by emphasizing the distinctive features. Emphasis may be achieved by exaggerating, caricaturizing, or otherwise enhancing them in a stimulus display. The processes underlying differentiation are said to be abstraction, filtering, and selective attention of sense organs. Effective training maximizes feature contrasts and eliminates nondistinctive features or noise.

The most original and interesting aspect of Gibson's specificity theory is the use of the model worked out by Roman Jakobson for the distinctive features of phonemes. This model enumerates a set of 12 features, each with a plus and a minus value forming a binary opposition, that suffices to give unique bundles for all the phonemes in all languages. The minimal distinction is one feature opposition between phonemes, but a pair may differ by a number of features. The confusability of a pair may be scaled by the number of distinctive features between them.

The model contains a built-in developmental hypothesis that is especially seductive. In speaking and comprehending speech, features develop by progressive differentiation. Evidence for a hierarchical pattern of development by successive binary fission is striking, and a systematic order of loss of features with aphasic dissolution shows the child's development in reverse. Gibson also tried out an analogous feature chart for graphemes with encouraging success. The features selected were those that electrophysiological studies showed to be critical: straight versus curved lines, and contrasting orientations. Other criteria were that the features be invariant under transformation of size, brightness, and perspective, be unique for each grapheme, and be reasonably economical.

The model is of particular interest because it is potentially applicable to complex educational problems—to the teaching of reading or languages, for example—as well as to more elementary discriminations of objects or forms.

But no physiological mechanisms are suggested and the problem of what neural basis might underlie the descriptive psychological processes is left untouched.

It is by no means clear why a comprehensive theory of perceptual learning and development must, in order to support a differentiation theory, cast a vote against association. As a psychological process with an underlying neural basis in the activity of the analyzers and the central nervous system, association is accepted by neurophysiologists. Once the discriminable aspects have been discovered in the stimulus array, the problem for the organism is how to combine several features into a single coding operation (as when, to borrow examples from Bower, we immediately judge a wine as dry or a woman as beautiful). Gibson does not face the question of how the pattern or the "unique bundle of features" gets connected with a particular differential response. Some process of association would seem to be necessary and would not invalidate the parts of the theory that account for dissociation or abstraction of new features. But Gibson says that "the attaching of the right response to the right cue is a trivial matter; it is discovering the critical differentiating feature that accounts for differences in difficulty . . ." (p. 332). This is a matter of learning to attend selectively to the critical variables.

But how, except in very general terms, does this increase in specificity occur? A wealth of evidence is presented to show that practice (repetition of stimulation) results in the reduction of errors, at least variable errors, and that the ability to detect regularity and structure grows with age. But the conditions under which practice results in learning, or age exerts its differentiating effects, are left vague or unspecified.

Is it not possible that the increased specificity of response that here provides the criterion of perceptual learning may equally well result from increased motivation or effort to judge? We can point to experimental evidence that variability of responding is a function of intention or attitude, sensitive to the control of instructions to 'be alert' or to 'pay attention' to particular stimulus aspects. The developmental trends that Gibson stresses may similarly reflect not perceptual modifications but increases in the efficiency of stimulus control over differential responses. Her own experiments fail to show critical ages for the appearance of attentional strategies. It may be that age is an incidental variable and that what appear to be developmental trends actually reflect the operation of stimulus-response contingencies. Perhaps, too, there are limits to the number of feature differences that ought to be maximized for optimal discrimination, limits that vary with the age or intelligence of the observer or with the difficulty of the task.

Who will read this book? The author hopes that in addition to colleagues interested in perceptual learning and development, students will find it useful, for advanced undergraduate courses in perception, child development, and education. It may be instructive to consider the book itself as an example of coded stimuli and to ask whether differentiation theory can be applied to account for its effectiveness as a textbook for the student of perceptual learning. How did Gibson herself apply the principles she urges as important when she set about to enhance the distinctiveness of her own theory?

After giving it a try in my course in perception, some students reported that they found the book difficult. Many favorable comments remarked the excellence of the summaries, the clarity of the organization, the well-marked sections following a logical order. The better the students, the more adept they were at discovering the distinctive features the author had provided, and they recognized the contrast between her theory and those against which it is opposed. As the book points out, an important predictor of discriminatory performance is the familiarity of features. New stimuli belonging to a class of objects having familiar distinctive features are more readily discriminated. Thus students with more background in psychology were able to take advantage of many subtle strengths of the text not available to the less sophisticated students.

As for her colleagues, there is no question but that the book will be widely read, studied, and discussed. It will serve as an important landmark and major reference work in the psychology of perception for years to come.

Alberta Gilinsky, *University of Bridgeport*

Learning, Memory, and Conceptual Processes

By Walter Kintsch. New York: Wiley, 1970. Pp. xii, 497. $9.95.

A good textbook should meet two criteria: (*a*) it should reflect accurately and in as up-to-date a way as possible the current state of the subject matter; (*b*) it should show the way to future development. The present text serves admirably on the first count but less well on the second. That this is the case should be a surprise to no one, for very few textbooks in psychology during the past half-century have played this important role. However, it does point up the fact that Kintsch's book accurately reflects the state of the study of human learning and memory. Therefore, many of the critical remarks that follow are not so much addressed to the book as to the field.

Granted that Kintsch's book is a faithful mirror of the current patterns of investigation in human learning and memory, one's chief disappointment in it is that it mirrors just as faithfully the disarticulation that exists in the field. There is no apparent effect in the excellent and informative chapters on memory of the research that is reported in the later chapter on linguistic organization (grammar). Worse still, the chapter on concept identification and the chapter on rule learning and language are only imperfectly articulated. This is disappointing because the bridge exists in the research on concept learning itself. Many of the important recent studies by Bourne and his group are assigned not to the chapter on concept identification but to the chapter on rule learning. There is nothing wrong with such a decision in itself, except that the author fails to tell us why there is a special chapter on concept identification which is apparently so independent of the material on rule learning. One could guess that the reason has to do with the fact that many of the theories faithfully summarized in the chapter on concept identification seem to ignore the logical structure of concepts in an effort to treat concept learning as a special case of associative learning.

The book is almost theoretically neutral. Almost, but not quite, for the stimulus-response tradition, the behavioristically oriented mathematical models of the 1950s and early 1960s receive the bulk of attention. A curious example of this prejudice is the chapter on discrimination learning. It is curious in the sense that so much of the material was obtained not from studies on human learning but from studies of animal learning under the implicit notion—guiding so much research in learning from 1930 through 1960—that the species' differences, including those between human beings and other species, are unimportant. This lends to the early material in this chapter an old-fashioned flavor, an impression reinforced by the date of some of the citations. To be applauded is the introduction of some discussion of judgmental material into this chapter, and the body of the chapter contains some interesting and important work on speech discrimination that ordinarily would not get into a textbook devoted to human learning and memory.

One can carp, but in summary it should be said that Kintsch's book is almost the ideal textbook on the first criterion mentioned at the beginning of this review. That so faithful a reproduction is not as creative as it might be should scarcely be a cause for disappointment. In fact, when one remembers the enormous impact of Woodworth's 1938 text (which I still keep on my 'most active' bookshelf), the importance of accurate and faithful reporting is apparent.

James Deese, *Johns Hopkins University*

Verbal Learning and Memory: Selective Readings

Edited by Leo Postman and Geoffrey Keppel. Baltimore: Penguin Books, 1969. Pp. 501. $2.95.

Postman and Keppel have produced a catholic reading text in the traditional fields of verbal learning. A large section of the book is given over to the problem of acquisition, a problem that is not really at the forefront of research in verbal learning (which tends to be dominated by studies in interference and organization in memory). This heavy influence on acquisition is all the more curious in that there is no representation of the psycholinguistic studies. These have revived some interest in mechanisms of acquisition. Even more curious is the absence of studies of concept identification (particularly those concerned with rule learning) in this section. The rest of the book is excellent. There is a good section on organization and recall, a section which suffers only from the lack of one or more of the important papers by George Mandler. There is a surprisingly small section devoted to interference, but a good one. In summary, this book can be regarded as the faithful mirror of Berkeley-style verbal learning as of 1969.

James Deese, *Johns Hopkins University*

Molecular Approaches to Learning and Memory

Edited by William Byrne. New York: Academic Press, 1970. Pp. xxiii, 358.

This is a collection of papers delivered at a symposium on memory transfer, at the meetings of the American Association for the Advancement of Science, New York, December, 1967. The volume also includes a list of publications on the memory-transfer phenomenon up to October, 1969. In spite of the title of the symposium, there are also papers by Jarvik ("The Role of

Consolidation in Memory"), Barondes ("Inhibitors of Protein Synthesis and Memory"), Agranoff ("Memory Formation in the Goldfish"), Zemp et al. ("The Effect of Training on the Incorporation of Radioactive Precursors into RNA of Brain"), Bennet et al. ("Differential Experience and Brain Measures"), Appel ("Macromolecular Synthesis in Synaptosomes"), and Essman ("Purine Metabolites in Memory Consolidation"). There is also a short communication by Melvin J. Cohen on the "Cytological Correlates of Functional Change in Central Neurons of Insects" and a valuable paper by Corning and Riccio on the "Planarian Controversy."

The short and lucid accounts of the work of these investigators will be useful to students, especially as some of the discussion following the papers has been incorporated. The rest of the volume, apart from an excellent summary by Roy John, is devoted to the controversial problem of memory transfer by means of brain extracts. It has become increasingly clear that the behavior of recipient animals can be differentially affected by some aspect of the previous treatment of the donor animal. What is yet unclear is whether the substrate of learned behavior is transferred. Unfortunately, the experimental design to clarify this must be quite complicated. Many workers do not seem to be aware that even within a discrimination task, the recipient animal's choice need not be determined by some cognitive factor but simply by some motivational alteration. A striking example of this was found in the reviewer's laboratory. It is believed that goldfish have a preference for red, learning to go to red somewhat more easily. This is true when they are hungry. But if they are given a choice between blue and red when escaping from shock, they choose blue. They cannot learn to swim to red to escape from shock and persist in swimming to blue even though shock persists. Similar though less striking alteration of biases can be observed in rats. Although there is considerable scepticism about the existence of memory transfer, the present volume does not reflect its presence or air the legitimate doubts on the issue. The book would be better if it were not so one-sided.

J. Anthony Deutsch, *University of California at San Diego*

Basic Color Terms: Their Universality and Evolution

By Brent Berlin and Paul Kay. Berkeley and Los Angeles: University of California Press, 1969. Pp. xii, 178. $8.50.

The authors are anthropologists who have surveyed in a simplistic manner the evolution of 11 color names (not 'terms' in the usual sense) in 20 different language groups. The color names with which they were concerned are white, black, red, yellow, green, blue, brown, pink, purple, orange, and grey (*sic*). The language groups were Arabic, Bulgarian, Catalan, Cantonese, Mandarin, English, Hebrew, Hungarian, Ibibio, Indonesian, Japanese, Korean, Pomo, Spanish, Swahili, Tagalog, Thai, Tzeltal, Urdu, and Vietnamese.

The study, though interesting, will have most appeal to psychologists concerned with the evolution of linguistics, but least pertinence for those concerned with the perceptual aspects of color.

Robert W. Burnham, *Eastman Kodak Company, Rochester, New York*

Light, Color, and Environment

By Faber Birren. New York: Van Nostrand Reinhold, 1969. Pp. 131. $19.50.

Though heralded on the title page as "a thorough presentation of facts on the biological and psychological effects of color—plus historical data and detailed recommendations for the resultful use of color in modern human environments," this book's scientific base is thinly conceived, and the recommended applications of color will be of more interest to designers and decorators than to experimentalists concerned with the psychophysiological aspects of color.

Robert W. Burnham, *Eastman Kodak Company, Rochester, New York*

Olfaction and Odours: An Osphrésiological Essay

By William McCartney. Berlin: Springer Verlag, 1968. Pp. viii, 249. $9.00.

The author of this book states in his foreword that "no truly comprehensive modern monograph on olfaction exists" (p. iii). Unfortunately, those individuals who are either casually or actively interested in olfaction and would like to see this void filled must remain disappoined, as McCartney goes on to note that his book "also is not comprehensive" (p. iv).

A sampling of the topics touched upon in this volume includes: an assessment of the olfactory prowess of dogs, birds, and fish; a comparison of olfaction in blind versus sighted and civilized versus savage men; a survey of the various odor-classification systems; some of the proposed relationships between the molecular structure of a compound and its odor; and several of the theories of olfactory stimulation. Although the author acknowledges that several of these topics are central issues in modern olfaction and have broad practical and theoretical import, the amount of space devoted to each topic in no way reflects its relative importance. For instance, nearly one-third of the text is devoted to a review of the odor-tracking ability of the dog, with particular emphasis on reports contained in the older and more obscure popular and scientific literature. On the other hand, the anatomy of the olfactory system is treated in one paragraph and the very important topic of odor communication among individuals by pheromones is not treated at all.

This book could prove useful to those interested in the historical background and development of olfaction. Individuals seeking information about current ideas and problem areas in modern olfaction should look elsewhere.

Robert J. O'Connell, *Rockefeller University*

Foundations of Abnormal Psychology

Edited by Perry London and David Rosenhan. New York: Holt, Rinehart and Winston, 1969. Pp. x, 644. $9.95.

This book represents a departure from the traditional format of a text in abnormal psychology, and the editors' desire to create a dialogue between general and abnormal psychology is admirable. Taken individually, some of the chapters represent excellent coverage of such diverse topics as behavior genetics, genius, behavior dynamics, social factors in psychopathology, and so on; many of these chapters would make good supplementary readings.

The primary weakness of this book as a text is that the individual contributions are never integrated and that the chapters dealing with descriptive psychopathology and treatment offer little more than brief versions of standard approaches to the area. The atmosphere of a dialogue is never created, perhaps because the editors did not include enough of the contributions of general psychology. The marriage between general and abnormal psychology would have been more compatible if the editors had included more extended treatment of areas such as learning, perception, and motivation in lieu of chapters on character, genius, and creativity.

Harold J. Johnson, *Bowling Green University*

Robots, Men, and Minds

By Ludwig von Bertalanffy. New York: George Braziller, 1967. Pp. x, 150. $5.00.

In 1928 a review appeared of the previous half-century of embryology, "Modern Theories of Development" by L. von Bertalanffy. In this remarkable essay (translated into English in 1933, reprinted by Harper in 1962), Bertalanffy perceptively outlined the history of the controversy between the 'mechanists,' who viewed the development of the embryo as the sum of the separate developments of its parts, and the 'vitalists,' who viewed embryonic differentiation as organized by an overall plan or 'entelechy,' leaving every cell equipotential to play any parental role. Both points of view were correct in a sense: cells may start out largely equipotential, but they are directed toward particular internally-guided growth patterns by critical interactions with neighboring cells. It will be surprising if similar interactions are not the empirical resolution of current investigations of brain and behavior.

Bertalanffy and others argued, from the inadequacy of either model of embryonic growth to account fully for the facts, that one must adopt an 'organismic' view of embryonic development. On this view, there is a *system* that transcends individual cell-to-cell interactions—a system which may have laws of its own that do not necessarily apply to the parts but only to the whole.

This argument was, and is, not unfamiliar to biologists and psychologists, who recognize the futility of misplaced attempts at mechanistic reductionism. However, to postulate a theory of higher-level 'wholes' is not enough: one must also develop the 'whole theory' and isolate the laws that pertain to it alone. Bertalanffy proposes that such laws of higher-order 'systems' are *general* to living (and perhaps nonliving) systems, independent of their mode of empirical reflection. That is, the same laws might apply to social systems, the dynamic structure of freight yards, language behavior, and so on. Indeed, various attempts to analyze specific problems in an abstract manner have filled the pages of a journal, *General Systems*, started by von Bertalanffy and others in 1956. The most general argument consistent with such investigations is the concept of (meta-) theoretical sciences—theoretical biology, theoretical history, and so on—in which the first problem is to isolate the laws of what *could* happen rather than what does and has happened.

The difficulty is that from deep enough a well of abstraction *all* problems appear to be the same. Even worse, with a sufficiently unconstrained mathematical system, we can make them *look* the same regardless of otherwise-apparent profound differences. As Bertalanffy himself hints in his later works, such abstract investigations may reveal more about the way humans conceive of the world than about the world itself. But, of course, it is the tenet of modern cognitive psychology that there are more direct ways of discovering the structure of man's mind.

In his monograph *Robots, Men, and Minds,* which is based on the Heinz Werner lectures of 1966, Bertalanffy emphasizes his view that modern psychology is still caught in the traditional doldrums of mechanistic views of man; he rejects the mechanistic view with a wide-ranging attempt to put psychology into humanistic and systematic perspective. Many of his individual points are correct and pertinent, but many are directed at personal demons that are of no relevance for psychology today. His main argument is that a man is a man, not a complicated rat, and should not be studied by using analogies and theories based on studies of lower organisms. He takes great pains to document his claim that modern psychology is based on "the positivist-behavioristic-commercialistic philosophy deeply ingrained in American life and thinking" (p. 116). His most concerted attack is on instrumental conditioning as typical of modern psychology, although he himself claims to be distilling beliefs general to the field rather than attacking a particular formulation.

There are two flaws in his argument; first, he fails to appreciate the recent redirections away from reductionist mechanism in the study of the mind; second, he fails to recognize that the field of operant conditioning can be viewed as an empirical study of the controlling parameters in the generalized learning capacity of *any* complex organism that is catabolically sensitive to stimuli and that anabolically compounds responses. It is a tenet of this school of psychology that there are general 'systematic' laws governing the linking of stimulus configurations and response patterns independent of the particular animal or learning system. So it is puzzling that Bertalanffy does not hail the study of operant behavior as one of the great attempts to develop an empirical base for investigating the laws of learning *as a system,* independent of any particular organism or task.

Of course, it is the extent to which operant psychology is species- and task-free that has led many students of behavior to reject it as an account of species-specific behaviors such as language. But Bertalanffy does not recognize this rejection as an indication of a change of heart. He passes over the criticisms of instrumental conditioning that have appeared recently as "concessions hesitantly made under the impact of circumstances" rather than seeing them as significant indicators of a change of orientation by many researchers concerned with humans and animals. Nevertheless, since 1963 (the date of his latest referenced criticism of instrumental conditioning), there has been a general shift within such traditional experimental areas as developmental psychology, the psychology of language, and information processing, a shift toward wholistic, species-specific accounts of behavior patterns and capacities. If Bertalanffy is not attacking a straw man, he is at least attacking one made of plaster of paris.

Robots, Men, and Minds appears to have been written partially as an exercise in the application of 'systems theory' to the writing of books about psychology and man. Bertalanffy is definitely aiming at The Big Picture. The distance from the subject is great and the generalizations are sweeping. One can grasp an impression of the book's style simply by skimming the pages for references to other thinkers: the sheer number of such references is overwhelming and their range leaves one awash in summary judgments about complex and differing views. Consider, for example, the possible concatenations of ideas of and about these individuals: Skinner, Maloney, Aristotle, Hitler, and the Beatles (p. 19); Maritain, Luther, Saint Jerome, Jesus, Lorenz, and Kant (p. 43); Lorenz, Köhler, Freud, and Koestler (p. 36); Meerloo, Werner, Piaget, Schachtel, and Freud versus Uexkuell, Cassirer, Humboldt, Whorf, Cantril, and Heisenberg (pp. 92-93). Generalizations encompassing such a variety of thinkers are necessarily distant from the subject matter; indeed, they may tell the reader more about Bertalanffy than anything else. Perhaps because of that, the number of provocative insights and opinions about a wide range of topics in current science and modern life make this brief book stimulating for the reader who can differentiate relations from associations.

T. G. Bever, *Columbia University*

Psychological Science and the Educational Enterprise

By Burton G. Andreas. New York: Wiley, 1968. Pp. ix, 103.

This short book is described by the author as the "emboldened effort of a psychologist to teach teachers some lessons in learning." Ideas from the fields of motivation, perception, learning, and memory have been organized into seven broadly prescriptive chapters. The results are not solid enough to help teachers make important instructional decisions, nor is the book likely to satisfy psychologists who hope to find evidence for the practical usefulness of their work.

The failure of Andreas's earnest effort reflects in part the poverty of psychology as a scientific discipline. Concepts in learning and memory are tender indeed, and one cannot go far from the laboratory bench before generalization from experimental findings becomes metaphor. But the book may also have erred about what psychological information can really be used by educators. Andreas consistently overvalues psychological findings or theoretical concepts and neglects the contribution of psychological methods to the educational enterprise. For example, he characterizes programmed instruction as a triumph of reinforcement theory. But many workers in programmed instruction see reinforcement theory as quite irrelevant to their work. By contrast, they consider the trial-and-revision methods used in program development as a logical outgrowth of the experimental procedures of the learning laboratory.

The book cites only 48 references and is given to unsupported pronouncements. It is clearly written, but it lacks substance.

Ernst Z. Rothkopf, *Bell Telephone Laboratories, Murray Hill, New Jersey*

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Their listing does not preclude their later review, however.)

Allen, Gay Wilson. *William James*. Minneapolis: University of Minnesota Press, 1970. Pp. 48. $.95.

Allen, Robert M., and Cortazzo, Arnold D. *Psycho-Social and Educational Aspects and Problems of Mental Retardation*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xii, 123. $6.00.

Argyle, Michael. *Social Interaction*. New York: Atherton Press, 1970. Pp. 504. $11.50.

Atkinson, Richard C., and Wilson, H. A., eds. *Computer-Assisted Instruction: A Book of Readings*. New York: Academic Press, 1961. Pp. x, 362. $6.75.

Baird, John C. *Psychophysical Analysis of Visual Space*. New York: Pergamon Press, 1970. Pp. viii, 321. $10.50.

Barber, Theodore X. *Hypnosis: A Scientific Approach*. New York: Van Nostrand Reinhold, 1970. Pp. 282. Paperback, $2.95.

Bardill, Donald R., and Ryan, Francis. *Family Group Casework: A Casework Approach to Family Therapy*. Washington, D. C.: Catholic University Press of America, 1969. Pp. vi, 69. $1.75.

Barton, Richard F. *A Primer on Simulation and Gaming*. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 239. $8.95; paperback, $4.95.

Bednarik, Karl. *The Male in Crisis*. New York: Knopf, 1970. Pp. xi, 194. $5.95.

Berger, Milton M. *Videotape Techniques in Psychiatric Training and Treatment*. New York: Brunner/Mazel, 1970. Pp. xv, 301. $15.00.

Berko, F. G.; Berko, J. M.; and Thompson, S. C.; eds. *Management of Brain-Damaged Children: A Parents' and Teachers' Guide*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 73. $6.50.

Black, Bertram J. *Industrial Therapy for the Mentally Ill*. New York: Grune & Stratton, 1970. Pp. 190. $9.75.

Bowyer, Laura R. *The Lowenfeld World Technique: Studies in Personality*. New York: Pergamon Press, 1969. Pp. xiv, 223. $6.50; paperback, $4.00.

Bower, Gordon, and Spence, Janet, eds. *The Psychology of Learning and Motivation*. New York: Academic Press, 1970. Pp. xiii, 398. $14.00.

Bourne, Peter G. *Men, Stress, and Vietnam*. Boston: Little, Brown and Co., 1970. Pp. 248. $10.00.

Boyle, D. G. *A Student's Guide to Piaget*. New York: Pergamon Press, 1969. Pp. ix, 156. $4.75; paperback, $2.25.

Braden, William. *The Age of Aquarius: Technology and the Cultural Revolution*. Chicago: Quadrangle Books, 1970. Pp. 306. $7.95.

Byrne, William L., ed. *Molecular Approaches to Learning and Memory*. New York: Academic Press, 1970. Pp. xxiii, 358. $18.00.

Carter, Hugh, and Glick, Paul C. *Marriage and Divorce: A Social and Economic Study*. Cambridge, Mass.: Harvard University Press, 1970. Pp. xxix, 451. $8.50.

Chassequet-Smirgel, J. *Female Sexuality: New Psychoanalytic Views*. Ann Arbor: University of Michigan Press, 1970. Pp. 220. $8.95.

Christy, Teresa E. *Cornerstone for Nursing Education*. New York: Teachers College Press, 1961. Pp. xii, 123. $8.95; paperback, $5.95.

Clark, K. E., and Miller, G. A. *Psychology*. Englewood Cliffs, N. J.: Prentice-Hall. 1970. Pp. 146. $5.95; paperback, $1.95.

Cogan, Peter L. *The Rhythmic Cycles of Optimism and Pessimism*. New York: William Frederick Press, 1969. Pp. 54. $5.95.

Collins, Ralph T. *Occupational Psychiatry*. Boston: Little, Brown and Co., 1970. Pp. 426. $12.50.

Coombs, C. H.; Dawes, R. M.; and Tversky, A. *Mathematical Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 419. $10.95.

Cox, Dunaway Rachel. *Youth into Maturity: A Study of Men and Women in the First Ten Years after College*. New York: Mental Health Materials Center, 1970. Pp. ix, 347. $10.00.

Critchley, MacDonald. *The Dyslexic Child*. 3d ed. Springfield, Ill.: Charles C Thomas, 1970. Pp. xv, 138. $6.50.

Cronbach, Lee J. *Essentials of Psychological Testing*. 3d ed. New York: Harper & Row, 1970. Pp. xxvii, 752. $10.50.

D'Amato, M. R. *Experimental Psychology*. New York: McGraw-Hill, 1970. Pp. 727. $10.50.

Daniels, David N.; Gihula, Marshall F.; and Ochberg, Frank M.; eds. *Violence and the Struggle for Existence*. Boston: Little Brown and Co., 1970. Pp. 472. $12.50.

Danziger, K., ed. *Readings in Child Socialization*. New York: Pergamon Press, 1970. Pp. x, 337. $8.00; paperback, $5.50.

Davies, D. R., and Tune, G. S. *Human Vigilance Performance*. New York: American Elsevier, 1970. Pp. 291. $11.50.

De Groot, A. D. *Methodology*. New York: Humanities Press, 1970. Pp. xviii, 400. $12.50.

Dember, William N., and Jenkins, James J. *General Psychology: Modeling Behavior and Experience*. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 784. $9.95.

Dessart, Charles, ed. *Psychologie et sciences humaines*. Paris: SEDIM, 1970. Pp. 420. Fr 28.10.

Donchin, Emanuel, and Lindsley, D. B., eds. *Averaged Evoked Potentials*. Washington, D. C.: National Aeronautics and Space Administration. Pp. 374. $2.00.

Dreikurs, Rudolf, and Gray, Loren. *Parents' Guide to Child Discipline*. New York: Hawthorn Books, 1970. Pp. 101. $3.95.

Duster, Troy. *The Legislation of Morality.* New York: The Free Press, 1970. Pp. x, 274, $6.95.

Eck, Marcel. *Lies and Truth.* New York: Macmillan, 1970. Pp. 176. $5.95.

Elveton, R. O. *The Phenomenology of Husserl.* Chicago: Quadrangle Books, 1970. Pp. 306. $15.00; paperback, $2.95.

Emery, F. E. *Freedom and Justice within Walls: The Bristol Prison Experiment.* New York: Barnes & Noble, 1970. Pp. 110. $5.00.

Ertel, Suitbert. *Psychophonetik: Untersuchunger über Lautsymbolik und Motivation.* Göttingen: Verlag für Psychologie, 1970. Pp. 230.

Eysenck, Hans J., and Eysenck, Sybil B. *Personality Structure and Measurement.* San Diego, Calif.: Robert R. Knapp, 1969. Pp. xiii, 365. $8.95.

Fabun, Don. *Children of Change: On Motivation, You, and Creativity.* New York: Glencoe Press, 1969.

Fearing, Franklin. *Reflex Action.* Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 350. $3.45.

Feibleman, James K. *An Introduction to the Philosophy of Charles S. Peirce.* Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 503. $12.50; paperback, $2.95.

Fischer, William F. *Theories of Anxiety.* New York: Harper & Row, 1970. Pp. 174. $4.95.

Flanigan, P. J.; Baker, G. R.; and LaFollette, L. G. *An Orientation to Mental Retardation: A Programmed Text.* Springfield, Ill.: Charles C Thomas, 1970. Pp. 203. $9.50.

Freedman, Jonathan; Carlsmith, J. Merrill; and Sears, David O. *Social Psychology.* Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 493. $7.95.

Freeman, Lucy, ed. *Celebrities on the Couch: Personal Adventures in Psychoanalysis.* Los Angeles: Price/Stern/Sloan, 1970. Pp. 234. $5.95.

Fuller, Varden. *Rural Worker Adjustment to Urban Life.* Ann Arbor: University of Michigan Institute of Labor and Industrial Relations, 1970. Pp. 87. $2.25.

Furth, Hans G. *Piaget for Teachers.* Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 163. $6.95; paperback, $2.95.

Gabriel, John. *Children Growing Up.* New York: American Elsevier, 1970. Pp. 480. $6.75.

Galyin, Willard. *In the Service of Their Country: War Resisters in Prison.* New York: Viking Press, 1970. Pp. vi, 344. $6.95.

Gibbons, Don C. *Delinquent Behavior.* Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. ix, 276. $7.95.

Gibbs, C. A., ed. *Leadership.* Baltimore: Penguin Books, 1969. Pp. 439. $2.25.

Gilbert, G. M. *Personality Dynamics.* New York: Harper & Row, 1970. Pp. 353. $7.95.

Giorgi, Amedeo. *Psychology as a Human Science.* New York: Harper & Row, 1970. Pp. 240. $2.95.

Gladwin, Thomas. *East is a Big Bird.* Cambridge, Mass.: Harvard University Press, 1970. Pp. 241. $9.95.

Goethals, George W., and Klos, Dennis S. *Experiencing Youth: First-Person Accounts.* Boston: Little, Brown and Co., 1970. Pp. xiii, 399. $5.95.

Goethe, Johann Wolfgang. *Theory of Colours.* Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 423. $10.00; paperback, $2.95.

Gonzalez, Victor R. *Psychiatry and the Army Brat*. Springfield, Ill.: Charles C Thomas, 1970. Pp. ix, 87. $6.50.

Gordon, David C. *Self-love*. New York: Macmillan, 1970. Pp. 96. $4.95.

Grassi, Joseph R. *The Grassi Block Substitution Test for Measuring Organic Brain Pathology*. 2d ed. Springfield, Ill.: Charles C Thomas, 1970. Pp. 84. $6.50.

Grollman, Earl A. *Talking about Death*. Boston: Beacon Press, 1970. Pp. 29. $6.00.

Haas, Kurt. *Understanding Adjustment and Behavior*. 2d ed. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 519. $8.95.

Hall, Calvin S., and Lindzey, Gardner. *Theories of Personality*. 2d ed. New York: Wiley, 1970. Pp. xiv, 622. $9.95.

Hall, Calvin S., and Lund, Richard. *Dreams, Life, and Literature: A Study of Franz Kafka*. Chapel Hill: University of North Carolina Press, 1970. Pp. xii, 131. $6.00.

Halmos, Paul. *The Personal Service Society*. New York: Schocken Books, 1970. Pp. 214. $6.00.

Hartley, Ruth, and Goldeneon, Robert. *The Complete Book of Children's Play*. New York: Apollo Editions, 1970. Pp. 484. Paperback, $3.95.

Hartman, Sven S., ed. *Syncretism*. Stockholm: Almquist & Wiksell, 1970. Pp. 294.

Haskell, Martin R., and Yablonsky, Lewis. *Crime and Delinquency*. Chicago: Rand McNally, 1970. Pp. 528. $8.25.

Hatterer, Lawrence J. *Changing Homosexuality in the Male*. New York: McGraw-Hill, 1970. Pp. 492. $15.00.

Hayes, Nelson E., and Hayes, Tanya, eds. *Claude Levi-Strauss: The Anthropologist as Hero*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 264. $10.00; paperback, $2.95.

Hediger, H. *The Psychology and Behavior of Animals in Zoos and Circuses*, trans. Geoffrey Sircom. New York: Dover, 1969. Pp. vii, 166. $2.00.

Hensman, C. R. *From Gandhi to Guevara*. London: Pelican Press, 1969. Pp. 490. $8.95.

Herriot, Peter. *An Introduction to the Psychology of Language*. New York: Barnes & Noble, 1970. Pp. 168. $6.75.

Hershey, Gerald L., and Lugo, James O. *Living Psychology*. New York: Macmillan. 1970. Pp. 449. $7.95.

Hershey, Gerald L., and Lugo, James O. *Handbook for "Living Psychology."* New York: Macmillan, 1970. Pp. 271. $3.95.

Hiestand, Dale. *Discrimination in Employment*. Ann Arbor: University of Michigan Institute of Labor and Industrial Relations, 1970. Pp. 68. $2.00.

Hilgard, Joseph R. *Personality and Hypnosis*. Chicago: University of Chicago Press, 1970. Pp. x, 304. $12.50.

Hoffman, Adeline M. *The Daily Needs and Interests of Older People*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xxv, 493. $16.50.

Hogg, James, ed. *Psychology and the Visual Arts*. Baltimore: Penguin Books, 1969. Pp. 415. $2.45.

Holding, D. H. *Experimental Psychology in Industry*. Middlesex, England: Penguin Books, 1969. Pp. 428. $2.95.

Hoselitz, Bert F. *A Reader's Guide to the Social Sciences*. Rev. ed. New York: The Free Press, 1970. Pp. 425. $8.95.

Hutt, S. J., and Hutt, Corinne. *Direct Observation and Measurement of Behavior*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xii, 224. $14.95.

Jacobson, Edmund. *Modern Treatment of Tense Patients: Case Illustrations, Follow-ups, EMG Measurements*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xxviii, 452. $16.50.

Johnson, Marjorie Seddon, and Kress, R. A., eds. *Reading Difficulties: Classroom and Clinic*. Philadelphia: Temple University Press, 1970. Pp. 132. $4.00.

Kaufman, Harry. *Aggression and Altruism*. New York: Holt, Rinehart and Winston, 1970. $2.50.

Keen, Sam. *To a Dancing God*. New York: Harper & Row, 1970. Pp. xi, 160. $5.95; paperback, $1.95.

Kimble, Gregory A., and Garmezy, Norman. *Principles of General Psychology*. 3d ed. New York: Ronald Press, 1970. $9.50.

Kirk, G. S. *Myth*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 299. $7.95.

Kohler, Claude, and Bernard, Françoise. *Les Etats depressifs chez l'enfant*. Brussels: Charles Dessart, 1970. Pp. 210. Fr. 23.24.

Koster, W. G., ed. *Attention and Performance, II*. Amsterdam: North-Holland Publishing, 1969. Pp. 449. $19.60.

Kovel, Joel. *White Racism: A Psychohistory*. New York: Pantheon Books, 1970. Pp. xii, 300. $7.95.

Kriesberg, Louis. *Mothers in Poverty: A Study of Fatherless Families*. Chicago: Aldine, 1970. Pp. x, 356. $9.75.

Lamb, Warren, and Turner, David. *Management Behavior*. New York: International Universities Press, 1969. Pp. ix, 177. $6.50.

LeShan, Eda. *Natural Parenthood: Raising Your Child without a Script*. New York: Signet Books, 1970. $.75.

Levinson, Boris M. *Pet-Oriented Child Psychotherapy*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 202. $9.75.

Logan, Frank A. *Fundamentals of Learning and Motivation*. Dubuque, Iowa: Wm. C. Brown, 1970. Pp. xvii, 226. $3.95.

Lynd, Helen M. *On Shame and the Search for Identity*. New York: Harcourt, Brace & World, 1958. Pp. 318. $2.65.

McCulloch, Warren S. *Embodiments of Mind*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 402. $2.95.

McGlade, Francis S. *Adjustive Behavior and Safe Performance*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 193. $9.75.

MacKay, Donald M. *Information, Mechanism and Meaning*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 196. $8.95; paperback, $2.95.

Mackworth, Jane. *Vigilance and Habituation*. Baltimore: Penguin Books, 1969. Pp. 199. $1.95.

McNeil, Elton. *The Psychoses*. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 171. $6.95; paperback, $2.75.

Maher, B. *Adventures in Experimental Personality Research*. Vol. 5. New York: Academic Press, 1970. Pp. xii, 205. $9.50.

Marks, Isaac M. *Fears and Phobias*. New York: Academic Press, 1969. Pp. 302. $7.50.

Masters, William H, and Johnson, Virginia E. *Human Sexual Inadequacy*. Boston: Little, Brown and Co., 1970. Pp. 448. $12.50.

Mayer, Philip, ed. *Socialization: The Approach from Social Anthropology.* New York: Barnes & Noble, 1970. Pp. xxx, 352. $8.75.

Mehrabian, Albert. *Tactics of Social Influence.* Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 152. $6.95; paperback, $3.95.

Meldman, Monte Jay. *Diseases of Attention and Perception.* New York: Pergamon Press, 1970. Pp. xi, 243. $13.50.

Megargee, Edwin, and Hokanson, Jack. *The Dynamics of Aggression.* New York: Harper & Row, 1970. Pp. 271. $3.95.

Meltzoff, Julian, and Kornreich, Melvin. *Research in Psychotherapy.* New York: Atherton Press, 1970. Pp. 561. $15.00.

Mitscherlich, Alexander. *Society without the Father: A Contribution to Social Psychology.* New York: Schocken Books, 1970. Pp. 352. Paperback, $2.95.

Mittler, Peter, ed. *The Psychological Assessment of Mental and Physical Handicaps.* New York: Barnes & Noble, 1970. Pp. 857. $25.00.

Montague, Ashley. *The Direction of Human Development.* New York: Hawthorn Books, 1970. Pp. xii, 406. $7.95; paperback, $3.95.

Moray, Neville. *Listening and Attention.* Baltimore: Penguin Books, 1969. Pp. 107. $1.45.

Munro, Margaret. *Psychology and Education of the Young.* New York: American Elsevier, 1970. Pp. vii, 294. $4.95.

Mussen, Paul H.; Conger, John J.; and Kagan, Jerome; eds. *Readings in Child Development and Personality.* 2d ed. New York: Harper & Row, 1970. Pp. 595. $5.95.

Nash, John. *Developmental Psychology: A Psychobiological Approach.* New York: Prentice-Hall, 1970. Pp. xii, 583. $9.50.

Ornstein, Robert E. *On the Experience of Time.* Baltimore: Penguin Books, 1970. Pp. 125. $1.45.

Ostow, Mortimer. *The Psychology of Melancholy.* New York: Harper & Row, 1970. Pp. ix, 144. $5.95.

Palmer, James O. *The Psychological Assessment of Children.* New York: Wiley, 1970. Pp. 475. $12.50.

Pervin, L. A. *Personality.* New York: Wiley, 1970. Pp. xiv, 632. $9.95.

Pfeiffer, Carl, and Smythies, John. *International Review of Neurobiology.* Vol. 12. New York: Academic Press, 1970. Pp. 413. $21.00.

Piaget, Jean. *Genetic Epistemology,* trans. Eleanor Duckworth. New York: Columbia University Press, 1970. Pp. 84. $5.00.

Praeworsky, A., and Teune, H. *The Logic of Comparative Social Inquiry.* New York: Wiley, 1970. Pp. xiv, 153. $8.50.

Price-Williams, Douglass R. *Cross-Cultural Studies.* Baltimore: Penguin Books, Pp. 384. $2.75.

Pulton, E. C. *Environment and Human Efficiency.* Springfield, Ill.: Charles C Thomas, 1970. Pp. 328. $15.50.

Rapoport, Anatol. *N-Person Game Theory: Concepts and Applications.* Ann Arbor: University of Michigan Press, 1970. Pp. 331. $8.50; paperback, $3.95.

Rhine, Louisa E. *Mind over Matter: The Story of P. K.* New York: Macmillan, 1970. Pp. ix, 402. $7.95.

Rice, Joseph P. *The Gifted: Developing Total Talent.* Springfield, Ill.: Charles C Thomas, 1970. Pp. 339. $17.50.

Ricoeur, Paul. *Freud and Philosophy: An Essay on Interpretation*. New Haven, Conn.: Yale University Press, 1970. Pp. xii, 573. $15.00.

Roff, Merrill, and Ricks, David. *Life History Research in Psychology*. Minneapolis: University of Minnesota Press, 1970. Pp. vii, 321. $10.00.

Rogow, Arnold A. *The Psychiatrists*. New York: G. P. Putnam's Sons, 1970. Pp. 317. $7.95.

Rosenblueth, Arturo. *Mind and Brain: A Philosophy of Science*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. xii, 128. $5.95.

Rossman, Isadore. *Two Children by Choice*. New York: Parents' Magazine Press, 1970. Pp. 151. $3.95.

Rubin, Sol. *Crime and Juvenile Delinquency*. Dobbs Ferry, N. Y.: Ocean Publications, 1970. Pp. 234. $6.50.

Rubin, Vera, and Zavalloni, Marisa. *We Wish to be Looked Upon*. New York: Teachers College Press, 1969. Pp. xi, 257.

Sabin, Theodore, R., and Coe, William C. *The Student Psychologist's Handbook*. New York: Harper & Row, 1970. Pp. 115. $5.95; paperback, $2.95.

Sareyan, Alex. *The Psychiatrist and Public Issues*. New York: Group for the Advancement of Psychiatry, 1970. Pp. 32. $1.00.

Sareyan, Alex. *Psychotherapy and the Dual Research Tradition*. New York: Group for the Advancement of Psychiatry, 1970. Pp. 64. $1.50.

Sarris, Mary A., and Pennekamp, Marianne. *Collaboration in School Guidance*. New York: Brunner/Mazel, 1970. Pp. 243. $6.95.

Schultz, Duane P. *Psychology and Industry*. New York: Macmillan, 1970. Pp. x, 513. $5.95.

Scott, Ellis L., and Bolz, Roger W., eds. *Automation and Society*. Athens, Ga.: Center for the Study of Automation and Society, 1970. Pp. 214. $5.75.

Shausa, Anselm L., ed. *Where Medicine Fails*. Chicago: Aldine, 1970. Pp. vii, 171. $2.45.

Shiplett, C. Duane. *Put Some Stuff in Your Personality: An Action Course in Developing Self*. New York: Exposition Press, 1969. Pp. 103. $4.00.

Simkins, Lawrence D. *The Basis of Psychology as a Behavioral Science*. Waltham, Mass.: Blaisdell Publishing, 1969. Pp. 198. $3.25.

Skidelsky, Robert. *English Progressive Schools*. Baltimore: Penguin Books, 1969. Pp. 272. $1.45.

Slater, Philip. *The Pursuit of Loneliness: American Culture at the Breaking Point*. Boston: Beacon Press, 1970. Pp. xiii, 154. $7.50.

Smart, Frances. *Neurosis and Crime*. New York: Barnes & Noble, 1970. Pp. xx, 178. $6.75.

Smith, Milton G. *A Simplified Guide to Statistics for Psychology and Education*. 4th ed. New York: Holt, Rinehart and Winston, 1970. Pp. 244. Paperback, $3.95.

Sofer, Cyril. *Men in Mid-Career: A Study of British Managers and Technical Specialists*. New York: Cambridge University Press, 1970. Pp. xxii, 376. $11.50; paperback, $3.45.

Solomon, Arthur; Perry, Steven; and Devine, Robert. *Interpersonal Communication: A Cross-Disciplinary Approach*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 109. $8.50.

Spicker, Stuart F. *The Philosophy of the Body*. Chicago: Quadrangle Books, 1970. Pp. viii, 367. $15.00; paperback, $2.95.

Spielberger, Charles D.; Fox, Robert; and Masterton, Bruce. *Student Workbook for "Principles of General Psychology."* New York: Ronald Press, 1970.

Stein, Calvert. *Practical Pastoral Counseling*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 283. $10.75.

Szent-Gyorgyi, Albert. *The Crazy Ape*. New York: Philosophical Library, 1970. Pp. 93. $3.95.

Tharp, Roland, and Wetzel, Ralph. *Behavior Modification in the Natural Environment*. New York: Academic Press, 1970. Pp. 229. $10.00.

Thorndike, Edward L. *Human Nature and Social Order*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 373. $15.00.

Tucker, Irving F. *Adjustment: Models and Mechanisms*. New York: Academic Press, 1970. Pp. 489.

Tumin, Melvin M., ed. *Research Annual on Intergroup Relations*. Chicago: Quadrangle Books, 1970. Pp. x, 310. $10.00.

Van de Vate, Dwight, ed. *Persons, Privacy, and Feeling*. Memphis: Memphis State University Press, 1970. Pp. 142.

Von Haller Gilmer, B. *Psychology*. New York: Harper & Row, 1970. Pp. 512. $9.95.

Von Hilsheimer, George. *How to Live with Your Special Child*. Chicago: Quadrangle Books, 1970. Pp. 272. $10.00.

Von Mering, Otto, and Kasdan, Leonard, eds. *Anthropology and the Behavioral and Health Sciences*. Pittsburgh: University of Pittsburgh Press, 1970. Pp. xi, 340. $12.95.

Weiner, I. B. *Psychological Disturbance in Adolescence*. New York: Wiley, 1970. Pp. xiii, 400. $12.50.

Wettler, Manfred. *Syntakische Faktoren im verbalen Lernen*. Bern: Hans Huber, 1970. Pp. 151.

Whitehead, Anthony. *In the Service of Old Age*. Baltimore: Penguin Books, 1970. Pp. 160. $1.25.

Wikman, K. Robert V. *Lachesis and Nemesis: Four Chapters on the Human Condition in the Writings of Carl Linnaeus*. Stockholm: Almquist & Wiksell, 1970. Pp. 123.

Winnicott, D. W. *The Child, the Family, and the Outside World*. Baltimore: Penguin Books, 1969. Pp. 239. $12.50.

Wisdom, John. *Paradox and Discovery*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 166. $2.25.

Wolpe, Joseph. *The Practice of Behavior Therapy*. New York: Pergamon Press, 1969. Pp. x, 314. $8.95; paperback, $5.95.

Wortis, Joseph. *Mental Retardation*. Vol. 1. New York: Grune & Stratton, 1970. Pp. 321. $19.75.

Zanforin, Mario. *Inhibition of Responses to Light during Pre-Pupation Behavior in Larvae of the Fly Sarcophaga Barbata*. Florence: Sansoni, 1969.

Zifton, Betty Jean. *Return to Hiroshima*. New York: Atheneum Books, 1970. Pp. 90. $5.95.

now available in reprint

PSYCHOLOGICAL MONOGRAPHS. GENERAL AND APPLIED

(American Psychological Association)
Vols. 1-52. Princeton; etc. 1895-1940

Psychological Monographs makes available longer researches, treatises, or collections of laboratory studies which it was important to publish in their entirety and as they were received. The length of these studies usually excluded them from other journals, but the depth and detail of their findings make them essential for advanced study in psychology.

Publication of the *Monographs* began in 1895 as a collateral activity of the *Psychological Review* and continued as a major source of longer, more comprehensive articles until 1967, when the individual journals published by the American Psychological Association began to publish supplementary volumes of the *Monographs* in their own areas.

For a complete list of available titles, and their prices, please write to our Sales Department.

JOHNSON REPRINT CORPORATION
111 Fifth Avenue, New York, N.Y. 10003

NEW FROM ILLINOIS

LANGUAGE TRAINING IN EARLY CHILDHOOD EDUCATION

Edited by Celia Stendler Lavatelli

The papers in this volume present promising practices in language training derived from current psycholinguistic thought, especially as an aid to disadvantaged children.

Part I offers a critical examination of current work on the subject and analyzes some aspects of language to be stressed in training the child to reach the full potentiality of self-expression.

Part II deals with training procedures, including a successful program in infant training, a model of language teaching for disadvantaged Mexican-American children, and a report on language programs in British Infant Schools.

Part III evaluates some procedures in vogue today, describes a useful language comprehension test, and examines the use of elicited imitation as a tool in studying language development. 1971. 208 pages. Paper, $2.45.

Photo courtesy of N.Y. State College of Human Ecology.
Photographer: George Lavris

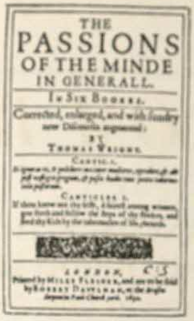
PSYCHOLINGUISTIC LEARNING DISABILITIES: DIAGNOSIS AND REMEDIATION

By Samuel A. Kirk and Winifred D. Kirk

This new monograph by the Kirks will help those using the Illinois Test of Psycholinguistic Abilities to interpret test results and organize remedial programs.

The book contains (1) a broad discussion of learning disabilities and an explanation of development and meaning of ITPA subtests; (2) a presentation of selected studies conducted with ITPA in which remedial teachers and psychologists will find research results of value in diagnosis and interpretation of subtest scores and patterns of disabilities; (3) general discussion of diagnosis followed by detailed description of sample patterns of disabilities and procedures by which these patterns are evolved; (4) suggestions for further analysis and diagnosis; (5) general guidelines for remediation; (6) specific suggestions from which teachers can formulate remedial procedures. 193 pages. Bibliography. Cloth, $7.95; paper, $2.45.

UNIVERSITY OF ILLINOIS PRESS Urbana · Chicago · London



THE
PASSIONS
OF THE MINDE
IN GENERALL.
In Six Bookes.
Corrected, enlarged, and with sundry
new Discourses augmented:
BY
Thomas Wright.

THE PASSIONS OF THE MINDE IN GENERALL

THOMAS WRIGHT

A Reprint Based on the 1604 Edition with an Introduction by Thomas O. Sloan

Wright's "pithie discourse about the passions of the minde" was one of the first originally English works on the physiological "humours" and the passions corresponding to them. Published in five editions between 1601 and 1630, his book has long been regarded as an important work in the history of psychology. Wright believed that studying the passions should contribute to man's efforts to lead a life of Christian virtue, but unlike other philosophers of his age, he included within his study of the passions lengthy discussions of rhetoric.

Sloan's introduction provides a brief biography of Wright and points out the importance of *The Passions* as a clear reflection of the cross-currents of political, religious, psychological, and rhetorical movements of its time. $12.95.

BIOLOGICAL AND SOCIAL FACTORS IN PSYCHOLINGUISTICS

EDITED BY JOHN MORTON

These nine essays discuss the extent of the uniqueness of human language and speculate on its origins. They present valuable new data from comparative study of the brain, the anatomy, and the behavior of living primates and of man.

The book is the result of a symposium held as part of the XIXth International Congress of Psychology in London during the summer of 1969. The contents are: "The Roots of Language" by Bernard Campbell, "Can Humans Talk?" by John C. Marshall, "Explaining Linguistic Universals" by David McNeill, "Comparing and Contrasting" by Roger Wales, "What Could Possibly Be Innate?" by John Morton, "The Grammar of Sign Language and the Problems of Language Universals" by I. M. Schlesinger, "Social and Natural in Semiotics" by A. A. Leontiev, "Syntax and Semantics" by Raqaiya Hasan, and "The Integrated Study of Language Behaviour" by Thomas G. Bever. $8.95.

UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

Touching: The Human Significance of the Skin

ASHLEY MONTAGU. A lively, absorbing, and pathbreaking inquiry into the importance of tactile experience in the development of the person. One turns from this book with new conceptions of right behavior in the family and much entertaining and unexpected knowledge of the constants and variations in human nurture around the world. $8.95

Perception of Other People

FRANZ FROM; translated by **BRENDAN A. MAHER** and **ERIK KVAN.** ". . . In this book From faces the question: How and under what conditions do we, literally, perceive the actions, the purposes and the properties of another person? The initial question is not that of veridicality: How correct are our judgments of other people?, but rather of the phenomenology of person-perception. From's studies represent a brilliant attempt to extend empirical-experimental psychology to include some of the less tangible phenomena of human experience."—Robert B. MacLeod $8.00

The Nonprofessional Revolution in Mental Health

FRANCINE SOBEY. The scarcity of factual information about paid and volunteer nonprofessionals across the country led the National Institute of Mental Health to sponsor the most comprehensive survey yet conducted, covering the work of some 10,000 people in over 185 government-sponsored mental health programs. The results of that survey, presented in this readable volume, provide an accurate picture of past experiences and current practices in the use of nonprofessionals in the field of mental health. $10.00

Columbia University Press

Address for orders: 136 So. Broadway, Irvington, N. Y. 10533
In Canada: McGill-Queen's University Press, 3458 Redpath St., Montreal 109

492/
23/3/7

Axc
Docket & the
issue to Librarian.
R.D.
13/3/72

Vol. 84, No. 3 Founded in 1887 by G. Stanley Hall September 19

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys **M. E. Bitterman**
Charles W. Eriksen **Benton J. Underwood**

E. L. J. LEEUWENBERG. A Perceptual Coding Language for Visual and Auditory Patterns 307

RUTH ELLEN GALPER and JULIAN HOCHBERG. Recognition Memory for Photographs of Faces 351

MAX COLTHEART and BARBARA ARTHUR. Tachistoscopic Hemifield Effects with Hemifield Report 355

ROBERT EFRON and DAVID N. LEE. The Visual Persistence of a Moving Stroboscopically Illuminated Object 365

DIRK C. PRATHER. Trial-and-Error versus Errorless Learning: Training, Transfer, and Stress 377

HOWARD R. MOSKOWITZ. The Sweetness and Pleasantness of Sugars 387

continued on back cover

Editors **Lloyd G. Humphreys** **M. E. Bitterman**
Charles W. Eriksen **Benton J. Underwood**
Editor Emeritus **Karl M. Dallenbach**
Book Review Editor **Julian Hochberg**
Consulting Editors **William Epstein** **Robert Fox** **Wendell Garner**
Lawrence E. Jones **John Krauskopf** **Allan Wagner**
Index Editor **Rand B. Evans** Necrology Editor **E. R. Hilgard**

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Eriksen (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Hochberg (Columbia University, New York, New York 10027).

Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text. Authors of articles may order offprints at the time they return galley proof; other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

The subscription price is $10.00 a year (foreign $10.50); single copies cost $2.75. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

© 1971 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

H. J. A. RIMOLDI, M. B. DONNELLY, and W. J. LAACK. Anchoring Effects in Psychological Scaling 407

SUSAN KARP MANNING. A Method to Assess Chance Alternation Levels in the Forced-Trial Paradigm 421

RICHARD L. GOTTWALD. Attribute-Response Correlations in Concept Attainment 425

ROBERT TEGHTSOONIAN and MARTHA TEGHTSOONIAN. The Apparent Lengths of Perimeters and Diameters Define a Ratio Smaller than Pi 437

The Sixty-seventh Annual Meeting of the Society of Experimental Psychologists (Carl P. Duncan) 439

The Sixty-third Annual Meeting of the Southern Society for Philosophy and Psychology (W. T. Blackstone) 440

The 1971 Arrowhead Conference (Benton J. Underwood) 442

Book Reviews (Earl Hunt, Herbert F. Crovitz, Stanley Coren, Donald C. Hood, W. B. Templeton, Ulric Neisser, David I. Mostofsky, Colin Blakemore, James R. Averill, Ian P. Howard; books received) 445

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS

Urbana, Illinois 61801

American Journal of Psychology
1971, Vol. 84, No. 3

A PERCEPTUAL CODING LANGUAGE FOR VISUAL AND AUDITORY PATTERNS

E. L. J. Leeuwenberg
University of Nijmegen, the Netherlands

A formal coding language for two- and three-dimensional shapes and for auditory patterns is presented. The coding allows the derivation of measures of information for both shapes and auditory patterns, as well as measures of preferred dimensionality, hierarchy, and substructural order for shapes—measures that rather successfully predict the corresponding perceptual phenomena. A theory of the origin of the units of information that are independently processed in the human perception of visual and auditory patterns is proposed.

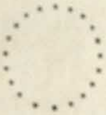
One of the most baffling functions of the human organism is undoubtedly its ability to represent the outside world. On thinking or deduction, so many 'laws' have been inferred (logic) that it has been seriously questioned whether man can in fact think at all. But on the induction of concepts, little or nothing is known.

The inductive process occurs even at the level of perception. Indeed, in perceiving two different triangles, it is immediately clear that the two shapes are in correspondence with respect to the category of triangularity. How does the perceptual system process the data of the patterns in such a manner that triangularity is selected as the relevant category, while in principle an infinite number of categories could be thought up? One might reply that the perceiver counts the number of angles of each figure and then checks to see whether both shapes have the same number of angles.

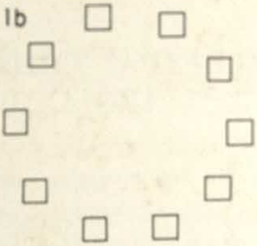
But what would the perceiver do with patterns la and lb (over)? Their common category is circularity. Clearly, counting the angles does not lead to this conclusion. Could it be that the perceptual system tests all possible (an infinite number of) rules in a fraction of a second? Or does it merely consider a small number of highly likely rules, which indirectly involve many others? We proceed from the latter suggestion and attempt to establish the existence of such universal coding principles.

Let us take a look at an example of how this investigation could be

1a

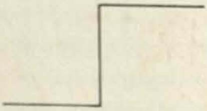


1b

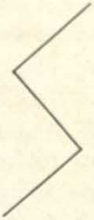


started. Assume that a perceptual system is presented with patterns 2a and 2b. Suppose too that we can take a look behind the eye of the perceiver and that we there discover how both figures are represented in his brain—or, equivalently, how they are expressed in a new and readable perceptual language. We may then read as the code for pattern 2a 'concatenated small straight lengths of line, which consecutively form the following angles in relation to a fixed base: 0 deg, 90 deg, 0 deg.' As the code for pattern 2b, we may read 'concatenated small straight lengths of line, which consecutively form the following angles in relation to a fixed base: 45, 135, 45 deg.'

2a



2b



The first parts of the two transcripts are the same, so we understand that the perceiver can indicate the correspondence of the straight lengths of line in the two patterns. But the last parts of the two codings do not show a ready-made similarity, so we know that the perceptual language we have used in our example is inadequate. Indeed, a human perceiver would see immediately that the patterns are near enough identical, the one merely being somewhat rotated in relation to the other. Let us therefore see if we can improve the language, can make it more *efficient*.

Suppose now that in the brain of the perceiver, there is still another perceiver (also known as homunculus) who is busy with the as yet irreconcilable values 0, 90, 0 deg and 45, 135, 45 deg—pairing them consecutively and subtracting one from the other. Thus 90 − 0 deg and 0 − 90 deg give 90 and −90 deg respectively, and 135 − 45 deg and 45 − 135 deg give 90 and −90 deg respectively. By means of this subtracting operation, the irreconcilable angle series are made equal. Thus the translation of a shape in terms of differences between consecutive angles is in this case a better one than the translation in terms of angles in relation to a fixed base.

For the time being, let us conclude from this only a rather general property. We saw in our second example that the perceiver first registered the angles. Only after this registration did the super-perceiver (homunculus) begin to perform his subtraction tricks—to build on the results of the first perceiver. Could it be that this strategy applies only in the translation of this particular pair of figures, or would it have a more general applicability?

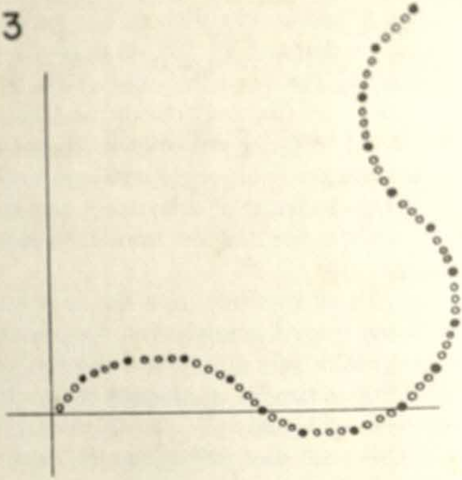
Let us apply the strategy to the translation of figure 1b. This could read 'dots in the configuration of a circle, in which a square is attached to each dot.' The first part of this description has to be determined first; only then can the second part specify the first. In other words, there is a *hierarchic* relationship between the two parts of the description. Homunculus's strategy has thus been applied in the translation of the figure. The translation, moreover, is psychologically relevant, for the first part of the translation describes exactly the 'psychological' similarity between figures 1a and 1b.

One might go on in this fashion, devising perceptual laws and testing them. But this will serve as a start. The way in which we have made this start may be very rough indeed; still, any perceptual rule conveying a broad restriction may be of use. Typical examples of such rules are provided by the Gestalt laws.

We now analyze a pattern in more detail in order to give a preliminary picture of the coding system we shall then discuss, and in order to illustrate more clearly the principle of perceptual hierarchy. Let us look again at how a pattern is stored by the perceiver—stored first at a peripheral level and then at a more central level, in such a manner that, in agreement with considerations by Attneave (1967), the representation at the latter level is less redundant than that at the more peripheral level.

For example, the peripheral representation of figure 3 might consist of

3



the series of angles formed at each dot, the angles between the dot's connections with the previous dot and the base axis. We then obtain, starting at the point at which the axes meet,

46, 46, 46, 46, 23, 23, 23, 23, 0, 0, 0, 0, −23, −23, −23, −23,
−46, −46, −46, −46, −23, −23, −23, −23, 0, 0, 0, 0,
23, 23, 23, 23, 46, 46, 46, 46, 69, 69, 69, 69, 92, 92, 92, 92,
115, 115, 115, 115, 138, 138, 138, 138, 115, 115, 115, 115,
92, 92, 92, 92, 69, 69, 69, 69, 46, 46, 46, 46.

Again we assume our homunculus, who is capable of translating this series of numbers into a shorter form. He does not, however, immediately cause the disappearance of the first redundant representation. First, the quadruplicity of the angle values gets expressed as

$$4 \cdot (46,23,0,-23,-46,-23,0,23,46,69,92,115,138,115,92,69,46).$$

Within the fences, homunculus's simplifications now begin to appear, if not the values themselves, but the differences between them are considered, as they were when homunculus performed his tricks on figures 2a and 2b. To see how this can be expressed symbolically, consider the minor example $0, 1, 2, 3, 2 = (1,1,1,-1)\int$. Within the fences preceding the $\int$ sign are the differences obtained pairwise. The last expression of figure 3 is thus retranslated as

$$4 \cdot ((46,-23,-23,-23,-23,23,23,23,23,23,23,23,23,$$
$$-23,-23,-23,-23)\int).$$

Once again, quadruplicity can be represented. It must be noted that this new representation is possible only after the translation by means of pairwise subtraction ($\int$) has been applied. Now, at the next recoding, a hierarchic transformation is performed. The translation now becomes

$$4 \cdot ((46,4 \cdot (-23,23,23,-23))\int).$$

Within the inner fences now lies a symmetry. This property constitutes, again in accordance with Attneave (1954), a simplifying perceptual principle.

Let us now agree that $R(a,b) = a, b, b, a$. Homunculus's recoding of figure 3 then gives

$$4 \cdot ((46,4 \cdot (R(-23,23)))\int),$$

which might also be written as

$$4 \cdot ((46,4 \cdot (R(\mp(23))))\int).$$

In this way, a figure is represented very concisely and very efficiently. One may attempt different sequences in the transformations, but, after some calculation, it will be apparent that only one sequence is the most efficient.

In this context, the very pertinent question is what determines the length or inefficiency of the formula. One might suggest that the greater the number of symbols in a formula, the more inefficient it is. For instance, in the formula $3 \times (a + b)$ seven symbols may be distinguished. However, these symbols are not independent of each other. The symbol $\times$ and the symbols ( ) have no meaning if 3 is absent, and vice versa. We therefore regard $3 \times (\;)$ as one given entity—as one information unit.

As a general rule, the $+$ and $\times$ symbols we call *operations*, while the ( ) symbols we call *indicators*. We attribute *information units* to the signs $\int$, R, $\pm$, and to numeric values. Thus $3 \times (a + b)$ contains but three information units. And the information in the last coding for figure 3 amounts to seven units. In general, we ignore the information concerning the tilt of the figure (in this case, 46 deg) and the first $\int$ transformation, so that the recoding for figure 3 contains five rather than seven units.

The coding system we are about to discuss has, in the first instance, bearing on two- and three-dimensional structures. It can also be applied, as will be shown, to auditory patterns. The coding system for two-dimensional figures has been described elsewhere (Leeuwenberg, 1968, 1969); we return to it frequently, because it contains simplifying analogies. Even so, our pattern language will give the impression of being very complicated at first, both because the analogies can become clear only afterward and also because we present the system of rules in a very formal fashion.

In presenting this model, we do not wish to suggest that there are no other possible codes for patterns, especially for the patterns in short-term memory or in imagination (Moran, 1969). We moreover leave out of consideration the manner in which representations of structures may in a certain context be laden with meaning when evoked in a human perceiver. In other words, the coding system we present is concerned with the ultimate cognitive representation of structures in long-term memory, insofar as these structures are free of meaning. From this coding system, such various perceptual properties can be deduced as the information, similarity, hierarchy, regularity, and redundancy of figures. By means of such quantifiable properties, quantitative tests of perception, of organic cerebral disorders, and of fluid intelligence (Cattell, 1968) may eventually be developed. Another effect of this coding system is that it may yield suggestions toward models of pattern recognition. In our discussion, an attempt will be made to substantiate the choice of the coding rules that are set out and then illustrated below. First, however, the following two main sections set out the coding rules and provide illustrative examples.

RULES FOR THE CODING

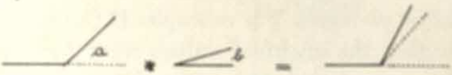
Information units

n Number: $a, a, a = 3 \cdot (a)$. This information may specify any size.
$\int$ Integration: $(3,2,5) \int = 0, 0 + 3, 0 + 3 + 2, 0 + 3 + 2 + 5 = 0, 3, 5, 10$.
R Reversal: $R\,\{3,2,5\} = \{3,2,5\}, \{3,2,5\}$.
$\pm$ Left-right variation: $\pm(90\text{ deg}) = 90\text{ deg}, -90\text{ deg}$.

Operations

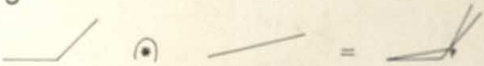
$+$ Addition: $5 + 3 = 8$.
$\times$ Multiplication: $5 \times 3 = 15$.
$\cdot$ Iteration: $3 \cdot (a) = a, a, a$.
$\circ$ Addition: This corresponds to the multiplication of vectors shown in figure 4. This operation does not affect length, however.

4



⊙ Addition of coincident angles, as in figure 5: $[|0,0,0,45,45,45|] \odot [|10,10,10,10,10,10|] = \begin{vmatrix}10,10,10,10,55,55\\ 0,\ 0,\ 0,45,45,45\end{vmatrix}$. This operation holds only

5



if two figures meet each other (point p). The operation is used in figure 5 in order to maintain certain patterns, like $|10,10,10,10,10,10|$, invariant under topological transformations such as the 'curved surface' $|\ 0,\ 0,\ 0,45,45,45|$. We have examples of this below, in the discussions of figures 10g, 10j, and 10l. Later it will also become clear why these numbers are usually enclosed in the absolute signs, $|\ |$.

Indicators

{ { Continuation: $\{ a \{ = a, a, a, a, \ldots$ This continuation holds until this sequence meets 'something' (e.g., a line or surface).

{ } Chunking: $2 \cdot \{a,b\} = \{a, b\}, \{a, b\}$. The sequence within these signs is treated as one unbroken unit in every context.

( ) Border: these signs are rather meaningless and often redundant. For example, $2 \cdot (a,b) = a, a, b, b$. The enclosed values do have a common fate, but function separately. Furthermore, these signs enable the applicability of the rule that in the reconstruction of a formula, the most deeply embedded values are solved first.

[ ] Breakdown into elements: $[\{a,b\}] = a, b$. These signs have the opposite effect of { }. However, this is only partly the case. They divide the enclosed group into subgroups by one step only. For example, $[\{\{a,b\}, \{c,d\}\}] = \{a,b\}, \{c,d\}$.

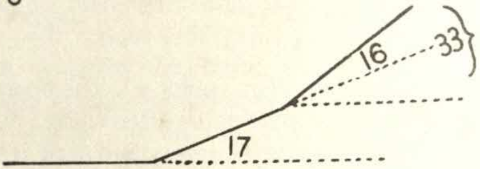
√ ∧ Reprisal: the function of an enclosed value continues until it is combined with all the unenclosed adjacent values. For example, $\surd 2, 3 \wedge \cdot (a,b) = 2 \cdot (a), 2 \cdot (b), 3 \cdot (a), 3 \cdot (b)$; whereas the normal situation is $(2,3) \cdot (a,b) = 2 \cdot (a), 3 \cdot (b)$.

[ ] Complete breakdown into elements: these signs represent the total

opposite of the { } signs. They break the strongest chunk into its utmost smallest elements. For example, $[\![\{\{a,b\},\{c,d\}\}]\!] = a,b,c,d$.

| | Absolute angles: the enclosed values refer to angles between sequentially connected straight lengths of line. These angles are, however, related to one base axis. When angles are *not* enclosed by these signs, then every value refers to the angle between a piece of line and the continuation of the previous line; these are 'relative angles.' For an example of absolute angles, |17 deg, 33 deg| = 17 deg, 16 deg (see figure 6).

6

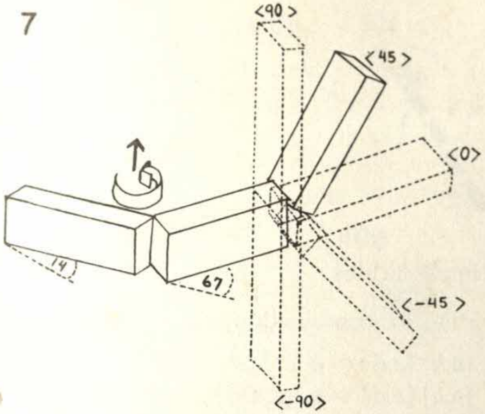


|| Independence of angles: $(a,b) \,||\, (c,d)$. This means that a and b do not form the reference axis for the angles c and d (in this case c and d are dependent on a and b in another aspect). Further, when a and b in their turn are dependent on p and v, and $||||$ is substituted for $||$, as, for example, in $((p,v)\,(a,b)) \,||||\, (c,d)$, then c and d are also angularly independent of p and v. There is an explanatory example in the next main section.

⟨ ⟩ 'Outerproduct': 14 deg, 67 deg ⟨45 deg⟩. The last angle, 45 deg, is in the direction of the outerproduct of 14 and 67 deg;[1] that is, in the direction given by the right-hand rule (see figure 7). In this case, 45 deg makes an outerproduct angle with 14 and 67 deg. This outerproduct enables us to describe three-dimensional figures in a way analogous to the way we describe two-dimensional figures, where we use the one past angle as reference base (axis) for the following angle. For three-dimensional figures, a surface determined by two angles (e.g., 14 deg, 67 deg) is used as reference base for every following angle.

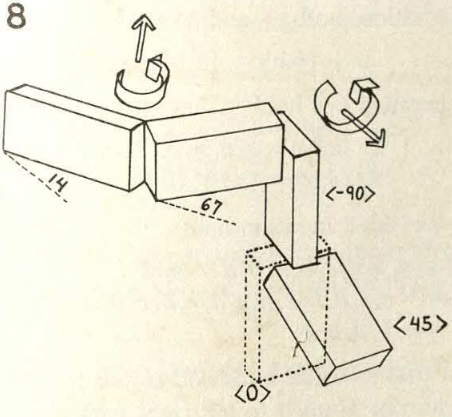
In cases (see figure 8) where there are more included angles, as 14 deg, 67 deg ⟨−90, 45 deg⟩, every included angle is also an outerproduct of the last two angles. Hence, 45 deg is an outerproduct of 67 and −90 deg. However, the right-hand rule must be applied to the rotation determined by 67, +90 deg. Note further that in the case of [14, 67 deg] $\overset{\frown}{*}$ [⟨|90, 45 deg|⟩], the 45 deg is not an

7



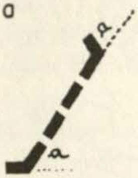
outerproduct of the last two angles, 67 and 90 deg, but of the paired angles at the left side of the ⊙ sign (in this case, 14 and 67 deg). This exception occurs only when the ⊙ operation is present. There is an explanatory example (see figure 10l) in the next main section.

8

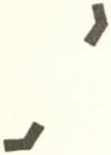


Vanishing sign: in figure 9a, $a, 2 \cdot (0), a, = a, 0, 0, a$; in figure 9b, $a, \overline{2 \cdot (0)}, a = a, \ , \ , a$. In the latter, the values function normally as existing elements but are invisible. An exception to this rule is the case in which there would be no difference between the situations with and without signs; then the values do not function normally (see the examples for figures 10c, 10d, and 10h in the next main section).

9a



9b



Combinatory and implicit rules

Combination: the rules for combination are

$$(a,b)\,(c,d) = a, c, b, d \quad \text{and}$$
$$(a,b)\,\{c,d\} = a, \{c,d\}, b, \{c,d\} \quad \text{and}$$
$$\{a,b\}\,(c,d) = \{a,b\}, c, \{a,b\}, d\,.$$

Operations: there are three rules that hold for all the operations (+, ×, ·, and *).

These are

$$(a,b) + (c,d) = a + c, b + d\cdot \quad \text{and}$$
$$(a,b) + \{c,d\} = a + \{c,d\}, b + \{c,d\} \quad \text{and}$$
$$\surd a,b\!/\!\!\wedge + (c,d) = \{a+c, a+d\}, \{b+c, b+d\}\,.$$

For addition (both + and *) holds

$$a + \{b,c\} = \{a+b,c\}\,.$$

For iteration (·) hold

$$2\cdot(a,b) = a, a, b, b \quad \text{and}$$
$$2\cdot\{a,b\} = \{a,b\}, \{a,b\}\,.$$

For one-sided iteration hold

$$3\,;(a,b)\,(c,d) = a, b, a, c, b, a, b, d \quad \text{and}$$
$$2\,{}^{:}(a,b)\,(c,d,e) = a, c, d, b, e, c, a, d, e, b, c, d, a, e, c, b, d, e \quad \text{and}$$
$$(1,2)\,{}^{:}(a)(b) = a, b, a, b, b\,.$$

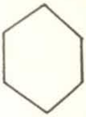
Note that the rules for iteration and for one-sided iteration can also be applied to left-right variation:

$$3\,;(\pm) = +, +, +, - \quad \text{and}$$
$$3\,{}^{:}(\pm) = +, -, -, - \quad \text{and}$$
$$3\cdot(\pm) = +, +, +, -, -, - \quad \text{and}$$
$$3\cdot\{\pm\} = \{+,-\}\,\{+,-\}\,\{+,-\}\,.$$

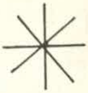
Serial structure: as in figure 10a,

$$a, 4\cdot(0), a, 4\cdot(0), \ldots = \{\, a,4\cdot(0)\, \{\,.$$

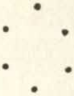
10a



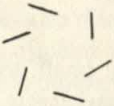
10b



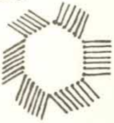
10c



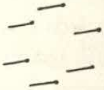
10d



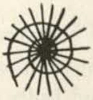
10e



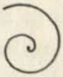
10f



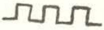
10g



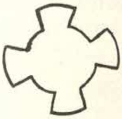
10h



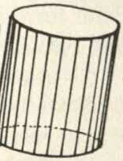
10i



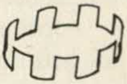
10l



10m



10n



Parallel structure: used at nodes, from which different branches derive, as in figure 10b (otherwise the normal serial structure is used).

$$\begin{matrix} a, 4 \cdot (0) \\ a, 4 \cdot (0) \\ a, 4 \cdot (0) \\ a, 4 \cdot (0) \end{matrix} = \overset{\sim\sim\sim}{\underset{\sim\sim\sim}{a,4\cdot(0)}}\,.$$

Note that the upper angles form the reference bases for the lower angles and have common starting points.

RULES: EXAMPLES AND ANALOGIES

Referring to some examples, we now attempt to show how visual structures of two and three dimensions, as well as auditory structures, may be coded with the aid of the rules just presented. At first, and generally, we shall express visual patterns in terms of *relative angles* (see the section on the absolute-angles indicator, | |, and figure 6 above). Let us now agree that the first 10 letters of the alphabet symbolize angles and that the remaining letters symbolize numbers.

Accordingly, figure 10a may be coded, as we saw above, as

$$\wr\, a,n\cdot(0)\, \wr\,.$$

If n is sufficiently small, this polygon approaches a circle. Thus, both shapes, polygon and circle, contain two information units. Further, as we also saw above, figure 10b may be described as

$$\overset{\sim\sim\sim\sim}{\underset{\sim\sim\sim\sim}{a,n\cdot(0)}}\,.$$

Now, just as the $\wr\ \wr$ symbols are used both in the serial formula for figure 10a and in the parallel formula for figure 10b, so may all combination rules be used both in the serial (horizontal) formulas that are normally appropriate to two-dimensional shapes and in the parallel (vertical) formulas that are particularly suitable to three-dimensional figures. The parallel structure of the formula for figure 10e is an example. Its code reads, because though two-dimensional, nodes are involved,

$$\begin{array}{c} \boxed{\wr\, a,n\cdot(0)\, \wr} \\[1em] \overbrace{\underbrace{b,p\cdot(0)}}\,. \end{array}$$

This means, take an element (angle) of the polygon $\wr\, a,n\cdot(0)\, \wr$ and attach to it a straight line of length $p,p\cdot(0)$, in such a way that the line is linked at an angle, b, to the element of the polygon. As the combination rule (given by the fences) dictates, one must now return to the next element (angle) of the polygon $\wr\, a,n\cdot(0)\, \wr$. This element should now be attached, not to the straight line $p\cdot(0)$ just mentioned, but, as the vertical structure of the formula indicates, to the first-mentioned element of the polygon. Subsequently attach to this, again, a straight line $p\cdot(0)$ at angle b; and so on in the same manner.

An example in which a vanishing sign finds expression is figure 10c, whose code reads

$$\wr\, a,\overline{n\cdot(0)}\, \wr\,.$$

And analogous to figure 10e, figure 10d may be written as

$$\begin{array}{c} \boxed{\wr\, a,\overline{n\cdot(0)}\, \wr} \\[1em] \overbrace{\underbrace{b,p\cdot(0)}}\,. \end{array}$$

This latter already constitutes an exceptional case, since not only $n\cdot(0)$

itself vanishes but also its further function, this being revealed by the fact that where $n\cdot(0)$ has disappeared, $p\cdot(0)$ has also disappeared. If this had not been so, the formula would have also applied to figure 10e.

Next, we may show the meaning of the independence-of-angles indicator, $||$. Figure 10f may be coded as

$$\begin{array}{c} \overbracket{\underbracket{\{\, a,\overline{n\cdot(0)}\, \{}} \\ \hline\hline \\ \overbrace{\underbrace{b,p\cdot(0)}} \end{array} .$$

Angle b is indeed linked to the elements (angles) of the polygon $\{a,\overline{n\cdot(0)}\}$; however, it is not dependent on the angles, a, but on the absolute base to which the a angles were referred in the first instance. In figure X of Table 1, the independence-of-angles indicator is used twice in order to show that cruciform patterns are angularly independent not only of the circular structures but also of the sinus superstructure.

The following examples deal with the very complicated use of the addition-of-coincident-angles operation, $\overset{\frown}{\ast}$. Here we also express visual patterns in *absolute angles*. The formula for figure 10g is made up of two parts: that for a star and that for a straight line. However, the line in the figure is not in fact straight, for at every point ($[\![\,]\!]\ [\![\,]\!]$) where it intersects ($\cap$) the radials of the star, a constant absolute angle ($|70\text{ deg}|$) is made with ($*$) these radials. In this way, the line is transformed into a spiral by virtue of its relation to the star. The representation of figure 10g is

$$[\![\widetilde{\underset{\sim}{a,n\cdot(0)}}]\!]\ \overset{\frown}{\ast}\ [\![|k\cdot(70)|]\!]$$

and the formula for figure 10h, where the star 'vanishes' from the spiral, is

$$[\![\overline{\widetilde{\underset{\sim}{a,n\cdot(0)}}}]\!]\ \overset{\frown}{\ast}\ [\![|k\cdot(70)|]\!] .$$

The reason the angles on the right-hand side of the formula are absolute is that they form a constant 70 deg with the absolute reference axes formed by the radials of the star described in the left-hand part of the formula (see figure 10b). In fact, the left-hand pattern must first be constructed before the right-hand pattern can be superimposed on it. Once the star pattern has transferred its directional function on the line $|k\cdot(70)|$, the star as such can vanish.

A similar process can be seen in the relation between figures 10i and 10j. Figure 10i's representation,

$$m \cdot [\,|(0,180)(90)|\,]\,,$$

can be recognized in the formula for figure 10j. This is in agreement with the fact that pattern 10i can easily be perceptually recognized in pattern 10j, which latter is represented by

$$\underset{\sim\sim\sim}{\overline{\widetilde{[a,\}\,0\,\{]}}}\ \overset{\frown}{\cdot}\ [\![\,\overline{t\cdot(0)},m\cdot[\,|(0,180)(90)|\,]\,]\!]\,.$$

The left-hand side again describes a star with infinitely long radials. Even though this star pattern vanishes, it still has its effect on the figure that corresponds to the right-hand part of the formula. This part first follows one of the radials of the star for a certain length, $t\cdot(0)$, which length later vanishes, and then the castellated structure $m\cdot[\,|0,90,180,90,0,90,180,90,0,\ldots|\,]$ connects on, in such a way that it is turned by the invisible star, which provides a new base plane.

The same procedure holds for the representation of pattern 10l, in which pattern 10i also occurs, invariant under the topological surface transformation described in pattern 10k and represented by

$$\begin{array}{c} \overline{\underline{\{\,a,n\cdot(0)\,\{}} \\ \circ \\ \underbrace{\overbrace{\langle 90\rangle,m\cdot(0)}} \end{array}\,.$$

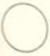
The representation of pattern 10l is

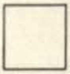
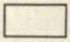
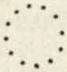
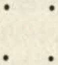
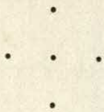
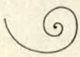
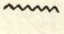
$$\left\|\ \begin{array}{c} \overline{\underline{\}\,\overline{a,n\cdot(0)}\,\}}} \\ \circ \\ \overline{\underbrace{\overbrace{\langle 90,0\rangle\,\}\,0\,\}}}} \end{array}\ \right\|\ \overset{\frown}{\cdot}\ [\![\,\langle m\cdot[\,|(0,180)(90)|\,]\rangle\,]\!]\,.$$

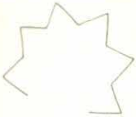
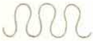
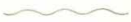
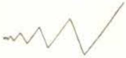
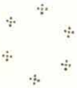
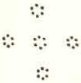
Pattern 10l is composed of patterns 10k and 10i. Pattern 10k is shown on the left side, and pattern 10i on the right side, of the formula. Because of the combination of the $\overset{\frown}{\cdot}$ and $\langle\ \rangle$ signs, the angles at the right-hand side form an outerproduct with the surfaces constituted by the angles of the formula at the left-hand side. However, this outerproduct occurs only between those right-hand angles that *coincide* with left-hand angles. In order to specify these angles, the left-hand figure has to be constructed first and then the right-hand figure must be superimposed on it step by step. So it is not possible to construct the shape of figure 10l directly from the two substructures, 10k and 10i.

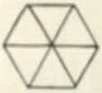
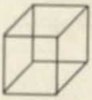
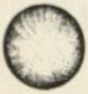
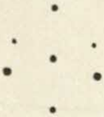
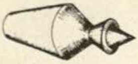
The proposed formulation of figure 10l is analogous to the construction of a drawing of figure 10i made on a curved surface. The usual way to construct figure 10l is to draw figure 10i on a sheet and then to bend the sheet in a circular way. The drawn figure and the bending of it are the constituents of the formula. This is a good instance of our general hypothesis (which is discussed more fully in the last section of this paper) that the subprocesses in the construction of a shape correspond to the units in the representation of the shape. The ⌢ operation also demonstrates the coding procedure for the formulation of two- and three-dimensional structures superimposed on three-dimensional shapes (see figure V in Table 1).

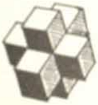
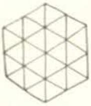
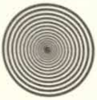
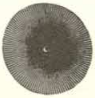
Table 1. Examples of the perceptual coding language applied to two- and three-dimensional shapes

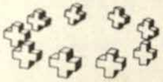
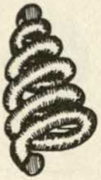
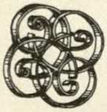
Shape	Formula	Information
A – 'grain'	(0)	0
B —— $n = 5$	$n\cdot(0)$ or $n\cdot A$	1
C . . . . . $n = 5$	$(n\cdot(0))\overline{\{n\cdot(0)\}}$ or $(B)\{\overline{B}\}$	2
D 	$\}a\}\{n\cdot(0)\}$ or $\}a\}\{B\}$	2

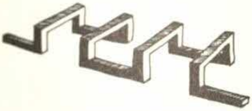
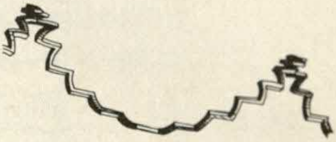
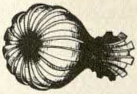
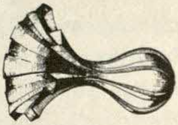
Shape	Formula	Information
E  $a = 90$ deg	same as for D $\}a\{\{n\cdot(0)\}$ or $\{a\{\{\mathrm{B}\}$	2
F 	$\}a\}(\{n\cdot(0)\},\{m\cdot(0)\})$ or $\{a\}(\{\mathrm{B}\},\{\mathrm{B}'\})$	3
G 	$\}a\}\{\overline{n\cdot(0)}\}$ or $\}a\}\{\overline{\mathrm{B}}\}$	2
H  $a = 90$ deg	same as for G $\}a\}\{\overline{n\cdot(0)}\}$ or $\{a\{\{\overline{\mathrm{B}}\}$	2
I	$\underset{\sim}{\tilde{a}}$ or D	1
J 	$\underset{\sim\sim\sim}{\overset{\sim\sim\sim}{a,\overline{n\cdot(0)},0}}$ or $\underset{\sim}{\overset{\sim}{a,\overline{\mathrm{B}}}}$	2
K 	$[m\cdot\{(a),\{n\cdot(0)\}\}]\int$ or $[m\cdot\{\mathrm{D}\}]\int$	4
L  $a = 90$ deg	$6\cdot\{\pm\}(a)$	3

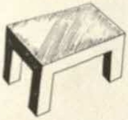
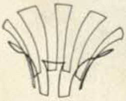
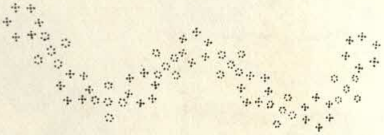
Shape	Formula	Information
M 	$((a)+[6\cdot\{\pm\}(90)])\{n\cdot(0)\}$ or $((a)+[\mathrm{L}])\{\mathrm{B}\}$	5
N  or 	$3\{(n\cdot(\pm)(a))\{n\cdot(0)\}\}$ or $3\{(n\cdot(\pm)\mathrm{D})\{\mathrm{B}\}\}$	5
O 	$\overline{n\cdot(0)},[6\{\pm\}(a)]\{\}0\}\}$ $\}\bar{0}\{$ $\overline{b,\}0\{}$ or $\bar{\mathrm{B}},[\mathrm{L}]\}\mathrm{A}\}$ $\}\bar{\mathrm{A}}\}$ $\overline{b,\}\mathrm{A}\}}$	5
P 	$\begin{array}{c}\{\overset{\frown}{\underset{\smile}{(a),\{\overline{n\cdot(0)}\}}}\{\\ \hline\hline \underbrace{\overbrace{\underset{\sim}{\overset{\sim}{90,\overline{m\cdot(0)},0}}}}\end{array}$ or $\begin{array}{c}\overset{\frown}{\underset{\smile}{\mathrm{G}}}\\ \overbrace{\underbrace{\mathrm{J}}}\end{array}$	4
Q 	$\begin{array}{c}\overset{\frown}{\underset{\smile}{\underset{\sim}{\overset{\sim}{90,\overline{m\cdot(0)},0}}}}\\ \}\overbrace{\underbrace{(a),\{\overline{n\cdot(0)}\}}}\}\end{array}$ or $\begin{array}{c}\overset{\frown}{\underset{\smile}{\mathrm{J}}}\\ \overbrace{\underbrace{\mathrm{G}}}\end{array}$	4

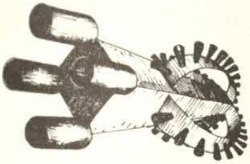
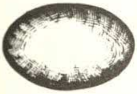
Shape	Formula	Information
R 	$\utilde{\widetilde{\{60,n\cdot(0)\{}}}$ or $\utilde{\widetilde{\mathrm{D}}}$	2
S 	$\{n\cdot(0),90\{$ $\langle 90\rangle,\{n\cdot(0),90\{$ or E $\langle 90\rangle\mathrm{E}$	4
S-1 	$\overset{\frown}{\underset{\smile}{\utilde{\widetilde{a}}}}$ $\circ$ $\overbrace{\underbrace{\langle 0\rangle,\{a,\{n\cdot(0)\}\{}}$ or $\overset{\frown}{\underset{\smile}{\mathrm{I}}}$ $\circ$ $\overbrace{\underbrace{\langle 0\rangle\mathrm{D}}}$	3
S-2  three-dimensional	$\utilde{\widetilde{\begin{array}{c}90,\overline{n\cdot(0)},0\\ \langle 90\rangle\overline{,n\cdot(0)},0\end{array}}}$ or J $\langle\mathrm{J}\rangle$	3
S-3 	$\overset{\frown}{\underset{\smile}{\mathrm{I}}}$ $\overbrace{\underbrace{\langle 60\rangle(\{n\cdot(0)\},\{m\cdot(0)\},\{p\cdot(0)\},\{q\cdot(0)\},\{r\cdot(0)\})(b,c,d,e)}}$ lengths angles	11

Shape	Formula	Information
S-4 	$\begin{array}{c}\frown\\ \sim\sim\sim\sim\\ 90,\overline{n\cdot(0)},0\\ \langle 90\rangle,\overline{n\cdot(0)},0\\ \sim\sim\sim\sim\\ \smile\\ =\\ \underbrace{\overbrace{\begin{array}{ll}\}\,n\cdot(0),90\,\{ & \\ & \langle 90\rangle\,\}\,n\cdot(0),90\,\{\end{array}}}\end{array}$ or $\begin{array}{c}\overset{\frown}{\underset{\smile}{\mathrm{J}}}\\ \langle \mathrm{J}\rangle\\ \overbrace{\underbrace{\mathrm{S}}}\end{array}$	7
S-5 	$\begin{array}{c}\sim\sim\sim\sim\sim\\ \begin{array}{ll}\}\,60,n\cdot(0)\,\} & \\ & 120,n\cdot(0)\end{array}\\ \sim\sim\sim\sim\sim\end{array}$ or $\begin{array}{c}\sim\sim\\ \begin{array}{ll}\mathrm{D} & \\ & \mathrm{D}'\end{array}\\ \sim\sim\end{array}$	3
T 	$\begin{array}{c}\sim\sim\sim\sim\sim\\ a,(n\cdot(0))\{\overline{m\cdot(0)}\}\\ \sim\sim\sim\sim\sim\end{array}$ or I,C	3
U 	$\begin{array}{c}\sim\sim\sim\\ a,n\cdot(0)\\ \sim\sim\sim\end{array}$ or I,B	2

Shape	Formula	Information
U-1 	$\overgroup{\underline{\underline{\undergroup{\{ a \{ \{\overline{n\cdot(0)}\}}}}}$ $\underbrace{\overbrace{\begin{array}{c}\boxed{\langle 90\rangle, \}\{m\cdot(0)\}[2;(\pm)(90)]\ \}} \\ \overbrace{\underbrace{\langle 90\rangle, k\cdot(0)}}\end{array}}}$	8
U-2 	$\boxed{[\![[m\cdot\{(a),\{n\cdot(0)\}\}]\int]\!]^*(\lvert\langle 45\rangle\rvert)}$ $\overbrace{\underbrace{90,\langle 0\rangle\ \{\ a,m\cdot(0)\ \{}}$ or $\begin{array}{c}[\![K]\!]^*(\lvert\langle 45\rangle\rvert) \\ \overbrace{\underbrace{90,\langle 0\rangle,D}}\end{array}$	8
U-3 	$\overgroup{\undergroup{\{\ 90,\{\overline{n\cdot(0)}\}\ \{}}$ $\underbrace{\overbrace{\begin{array}{c}\boxed{[m\cdot\{(a),\{p\cdot(0)\}\}]\int} \\ \overbrace{\underbrace{90,\langle 0\rangle,\ \}\ 90\ \}\ (\{q\cdot(0)\}\{r\cdot(0)\})}}\end{array}}}$ or $\begin{array}{c}\overgroup{\undergroup{H}} \\ \underbrace{\overbrace{\begin{array}{c}\overgroup{\undergroup{K}} \\ \overbrace{\underbrace{90,\langle 0\rangle,F}}\end{array}}}\end{array}$	10

Shape	Formula	Information
U-4 	$\overline{\overline{\underline{\underline{(\{n\cdot(0)\},\{m\cdot(0)\})[90,k\cdot\{(2,6)\cdot(\mp)(90,\langle 90\rangle)\}]}}}}$ $=$ $\overbrace{\underbrace{\wr\, s\cdot(0),90\,\wr \quad \langle 90\rangle,\wr\, s\cdot(0)90\,\wr}}$	13
U-5 	$\left[\dfrac{\begin{matrix}\overline{\overline{\underline{\underline{\wr\, a\,\wr\{m\cdot(0)\}}}}}\\ \ast\end{matrix}}{\overbrace{\underbrace{\langle 90\rangle,\wr\, 0\,\wr}}}\right]\circledast\overline{\overline{\underline{\underline{[\langle\|45,\{m\cdot(0)\}(s\cdot\{\pm\}(90))\|\rangle]}}}}$ $\overbrace{\underbrace{90,\langle 0\rangle,\wr\{k\cdot(0)\}(90)\wr}}$	11
	or $\left[\dfrac{\begin{matrix}\overline{\underline{\mathrm{D}}}\\ \ast\end{matrix}}{\overbrace{\underbrace{\langle 90\rangle,\wr\, 0\,\wr}}}\right]\circledast\overline{\overline{\underline{\underline{[\langle\|45,\{m\cdot(0)\}(\mathrm{L})\|\rangle]}}}}$ $\overbrace{\underbrace{90,\langle 0\rangle,\mathrm{E}}}$	
V  or 	$\left[\dfrac{\begin{matrix}\overline{\overline{\underline{\underline{\wr\, a,\{n\cdot(0)\}\,\wr}}}}\\ \ast\end{matrix}}{\{-90,\langle -b\rangle,\wr\, 0\,\wr\}}\right]\circledast\overline{\overline{\underline{\underline{\langle[m\cdot[\|(0,180)(90)\|]]\rangle}}}}$ $\ast \qquad \ast$ $\{-90,\wr\,((s,k)\cdot(\pm)\overline{\underline{(d)}})\{n\cdot(0)\}\,\wr\,\} =$ $\ast$ $\left[\!\left[\overbrace{\underbrace{90,\wr\, 0\,\wr}}\right]\!\right]\circledast\overbrace{\underbrace{[\wr\|-90\|\wr]}}$	15

Shape	Formula	Information
W 	$\{\,\overset{\frown}{(90)},\overline{\underline{(\{n\cdot(0)\},\{m\cdot(0)\})}}\,\{$ $\smile\quad\widetilde{\langle 90\rangle,s\cdot(0)}$ $\overset{\frown}{\underset{\smile}{\{\,90,k\cdot(0)\,\{}}}$ $\widetilde{\langle 90\rangle,r\cdot(0)}$	9
W-1 	$\left[\dfrac{\underset{\smile}{\{\,a,n\cdot(0)\,\}}\quad\circledast}{\widetilde{\langle 90\rangle^{\circledast}\langle 0\rangle,\{\,b,n\cdot(0)\,\}}}\right]\ \overset{\frown}{\circledast}\ [\langle\lvert(m,p)\cdot((0,180)(90))\rvert\rangle]$	9
X 	$\overline{\underline{2\cdot\{(k\cdot(\pm)(a))\{\overline{n\cdot(0)}\}\}}}$ $\left(\begin{array}{c}\overset{\frown}{\underset{\smile}{\{\,b\,\}\{\overline{s\cdot(0)}\}}}\\ \equiv\\ \widetilde{90,\overline{t\cdot(0)}}\end{array},\ \begin{array}{c}\overset{\frown}{\underset{\smile}{\widetilde{90,p\cdot\overline{(0)},0}}}\\ \\ \widetilde{\{\,b\,\}\{\overline{t\cdot(0)}\}}\end{array}\right)$ or	12 $\overline{\underline{N}}$ $\overset{\frown}{\underset{\smile}{\overset{\equiv}{P,Q}}}$

Shape	Formula	Information
Y 	$\wr 90 \qquad ,\{n\cdot(0)\}, \wr$ $\{^{*}\langle 90-f\rangle^{*}\langle f\rangle,p\cdot(0)\ \} \qquad \langle 0\rangle, \wr(k\cdot(a)(\{m\cdot(0)\})\wr$ $\{\langle -90,\tfrac{1}{2}n\cdot(0),\ \}$ $\overbrace{\overline{\overline{\underline{\underline{(90+f,-90)\{t\cdot(0)\}}}}}}$ $\overparen{\underparen{\langle 90\rangle,\langle 0\rangle}}$ $\underbrace{\overline{\overline{\underline{\underline{((w\cdot(1))\int);\wr(0)(a)\wr}}}}}$	$\{\underset{\sim}{\breve{b}},\langle -60\rangle,q\cdot(0)\}$ 20
Z 	$\overparen{\widetilde{\ }}$ a $\underparen{\widetilde{\ }}$ $\overbrace{\underbrace{\langle 0\rangle,((\wr a,n\cdot(0)\wr \not\int)\times 2)\not\int^{-1}}}$	4 $\not\int = \int \sin \int$ $\not\int^{-1} = \delta \arcsin \delta$

We have become acquainted with examples in which combination rules have been applied vertically. Now we introduce an example in which many combination rules (also vertically applicable) are applied horizontally. For a change, we employ an auditory pattern as our example. Thus the series of values shown in the first line of Table 2 corresponds not to angles but to a pattern of heights of tone. By applying the coding language, the values successively obtained lead to the coding in the last line of Table 2. This is the shortest description of the pattern—a pattern that continues much farther than is shown in the table, without any recurrence. The information units, shown by arrows, number 17.

The pattern of tones with which we have just dealt suggests an interesting analogy. As we have seen, the basic element for the description of two-dimensional shapes is the relative angle, which indicates how a

Table 2. An example of the perceptual coding language applied to a complex auditory pattern

Pattern

2,6,−2,7,3,6,−3,7,0,8,−1,6,2,7,−4,6,4,7,−5,8,5,6,−3,7,3,6,−3,7,4,8,−4,6,3,7,−2,6,1,7,−2,8,2,6,−3,7,3,6,
−0,7,4,8,−4,6,5,7,−5,6,0,7,−1,8,2,6,−3,7,3,6, . . .

Successive codings

} (2,−2,3,−3,0,−1,2,−4,4,−5,5,−3,3,−3,4, −4,3,−2,1,−2,2,−3,3,−0,4,−4,5−5,0,−1,2, . . .) (6,7,6,7,8) {
} (2,−2,3,−3,0,−1,2,−4,4,−5,5,−3,3,−3,4, −4,3,−2,1,−2,2,−3,3,−0,4,−4,5,−5,0,−1,2, . . .) [2·{6,7},8] {
} (±(2,2,3,3,0,1,2,4,4,5,5,3,3,3,4,4,3,2,1,2,2,3,3,0,4,4,5,5,0,1,2,3,3,4,4, . . .) [2·{6,7},8] {
} (±((3,1)⁚(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, . . .)(0,1,2,3,3,2,1,0))) [2·{6,7},8] {
} (±((3,1)⁚(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, . . .)[*R*{0,1,2,3}])) [2·{6,7},8] {
} (±((3,1)⁚(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, . . .)[*R*{(1,1,1)ʃ}])) [2·{6,7},8] {
} (±((3,1)⁚(2,2,3,3,4,4,5,5,3,3,4,4,2,2,3,3, . . .)[*R*{(3·(1))ʃ}])) [2·{6,7},8] }
} (±((3,1)⁚[√1,3,2/^+{1,1,2,2}][R{(3·(1))ʃ}])) [2·{6,7},8] }

Formula

} (±((3,1)⁚[√1,3,2/^+{2·(1,2)}][*R*{(3·(1))ʃ}])) [2·{6,7},8] }
↑ ↑↑ ↑↑↑ ↑ ↑↑ ↑ ↑ ↑ ↑ ↑ ↑↑ ↑
Information

straight line is determined only if the previous straight line is known or determined. Similarly, for three-dimensional shapes a straight line can be determined, via the outerproduct, only if the two preceding straight lines, whose sequential order is also given, are known or determined. It seems, therefore, that it is not absolute data (be they tones or lines), but rather the elementary internal relationships in a perceptual system, that constitute the basic data for coding. In the two sections that follow, we attempt to derive, and then to test, some measures of these relationships.

PROPERTIES OF THE CODING

In general, more than one formal representation can be found for one pattern. However, no cases have been found where more than one shortest, or most efficient, representation applies. Since the coding system is based on efficiency principles, the shortest formula (the one with the least information content) is chosen as the correct representation.

Information (I)

The information content can be determined by counting the number of information units that occur in the formula. The value 0, and therefore $\{0\}$ and $\langle 0 \rangle$, are not information. The value $n \cdot (0)$ conveys one unit of information. However, when $n \cdot (0)$ corresponds to length and has already appeared in the formula, then the other $n \cdot (0)$ does not convey information. There is an explanation of this farther below.

Every length or angle that differs from another in a shape corresponds to an information unit. The *precision* with which a length or angle is assimilated by a perceiver is beyond the scope of our theory, however. This type of dimensionless information about variation around a certain number can properly be estimated by the Shannon selective-information measure and is named the metrical information (M). Mackay (1950) distinguishes M from structural information (I), which latter corresponds to the number of dimensions or degrees of freedom with which the present study is solely concerned.

Although there is some tradeoff between I and M per dimension, there are indications that the transmitted M (total M) grows by increasing the dimensionality or logon content (Pollack and Klemmer, 1954). Whatever the interaction between M and I, the relative order of perceptual difficulty for different values of I is probably not disturbed by M. For example, the number of patterns regarded as identical or confused with each other increases with information content I (Handel and Garner, 1966). Let us assume that every structural information contains a certain amount of metrical information, which is not perceptually transmitted, and that the set of patterns indiscriminable from a given pattern with information I amounts to about $2^{I \times M}$. Thus, the higher the I, the larger the inferred set; the relative perceptual difficulty is unaltered although M and I may interact.[2]

As the surface raw data for the measurement of a given shape's I value, we use heights or absolute angles (angles related to a constant base axis) or relative angles (angles between successive straight lines). The latter type of data is more useful because the representation stays invariant under rotation of the pattern. In that way the coding is most isomorphic to the outside world, because the physical objects are in general not deformed in vision by a rotation of themselves or of the visual system. It is for this reason that we also assume that every angle is connected with the same ('grain,' or elementary) lengths. In this preference for relative angles, the coding system differs from the normal mathematical description of functions, in which the relationships between the projections of the figure are described with respect to an unchanging base axis and where it therefore makes sense to use length units on the base axis. For our purposes, however, absolute angles are useful raw data only when they produce a more economical description—when a figure is superimposed on a redundant curved surface, for example. Heights are used as raw data in analogous cases for similar reasons.

Changing from one level to another level of starting values without the change being reversed corresponds to one information unit. For the change from relative angles to absolute angles, one information transformation, $\int$, is required. Also, the step from absolute angles to heights corresponds to the information step $\sin \int$.[3] A special case occurs in a part of the representation of figure Z in Table 1, in which a circle is transformed to an ellipse. For reasons of economy, the circle is expressed in relative angles. By two steps the relative angles are changed to heights. These are enhanced with a certain factor (2 in this case), so that the circle is transformed to an ellipse. Then the heights are changed backward by the reverse of the same two steps, so that the ellipse is again expressed in the original relative angles. These two inverse transformations annihilate, however, the first two forward transformations and therefore annihilate also the corresponding units of information.

Another example is the transformation series $\int n \int^{-1}$. The first $\int$ changes the relative angles to absolute angles. The n stretches all grain lengths into $n \times$ 'grain' lengths of the sides of the patterns. The $\int^{-1}$ annihilates the otherwise-counted information, which corresponds to $\int$.

For this reason, only one information corresponds to the lengths of the sides when they are the same. Therefore, no information corresponds to the length coding $n \cdot (0)$ when this value has already been introduced in the formula.

The question now arises as to why the integrated relative angles are not used as starting values. In other words, why does a step from heights to absolute angles or from absolute angles to relative angles differ from the step from relative angles to integrated relative angles? The answer is that the first two steps correspond to a displacement and a rotation respectively, transformations under which the representation of patterns is invariant, whereas the step from relative angles to integrated relative angles corresponds to a transformation under which invariance of form is *not* preserved. For example, consider figure K in Table 1, a figure which can be efficiently expressed by integrated relative angles. This figure can be regarded as a circular wire, successively deformed by increasing forces. This transformation corresponds to a real physical deformation, while a displacement or a rotation usually keeps the physical qualities of an object constant. Since—and because of the way in which—our theory assumes that the perceptual laws are derived from physical laws, an information unit has to be assigned to this transformation.

This last illustration can be clarified, although the theory underlying it must wait until the last section of the paper. Suppose that the sensorimotor system makes a rotating movement and suppose that this movement is adjusted so that the final movement results in figure K. Then this movement corresponds to the scanning movement of figure K. However, the adjusting movement from circle to figure K evokes in the sensorimotor system the same deformative forces that are necessary to construct a wire shaped like figure K from a circular shape. The forces that are involved in the construction of this physical object are thus the very forces that are evoked in the sensorimotor system; and one information unit is assigned to the transformation (see Table 1, figure K, where the integration transformation applied to relative angles corresponds to that one information unit).

Similarity (S)

Another quantity that can be derived from the coding system—a quantity very directly related to information—is the similarity or difference between patterns. The difference, which we here call *d*, can be found by counting the number of necessary transformations applied to one figure in order to produce another figure. This d is used as a measure of the error in copy tests. When one pattern is not derived from another, a convenient symmetric measure is the relative difference,

$$D = \frac{d}{I_1 + I_2 - s}\,.$$

Here I_1 and I_2 are the information content of the two figures. The symbol s here refers to the number of common information units.

A definition can be proposed, namely, that relative similarity

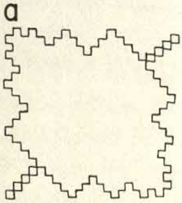
$$S = \frac{s}{I_1 + I_2 - s}\,.$$

It is reasonable to assume that all the possible information units inside the two patterns together consist of only d or s type transformations. Therefore, $I_1 + I_2 - s = d + s$ and $D + S = 1$. Probably it is easiest to estimate d, I_1, and I_2, from which D can be derived and then S from $S = I - D$.

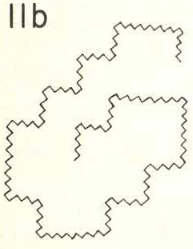
Hierarchy (H)

The drawback of the S measure is that it ignores the kind of differences between information units and their relations to each other, a drawback which limits the measure's meaningful applicability. Some transformations, however, considerably reduce the length of a formula; when they are not used in the representation of a pattern, the length of the formula is very much greater than when they are. And these transformations seem to be hierarchically related to their arguments. Defining hierarchy as *irreversible dependency*, we can then specify its amount by the information minus the number of unsubordinate relations. Thus figures 11a and 11b contain the same information ($I = 12$ units) but differ with respect to hierarchy ($H = 10$ and 6 units respectively). That is, for figure 11a the representation is

11a



11b



$$(\{\{R\{((2\cdot((3\cdot(1))\int))^{:}(0,2)(1))^{:}(X)\}\}(X)\})(n\cdot(0))$$
$$\uparrow$$

$X = 90, -90$ $\qquad\qquad$ $H = 12 - 2 = 10$
$\uparrow$

whereas for figure 11b it is

$$((\sqrt{}2,1,3,2,4,1/^{\wedge}\cdot\{R\{3\cdot\{90,-90\}\}\})(90))(n\cdot(0))$$
$$\uparrow\uparrow\uparrow\uparrow\uparrow \qquad\qquad \uparrow$$

$H = 12 - 6 = 6$

Hierarchy plays a very important role in the perception of figures. The theoretical weakness of this H quantity as a measure of regularity is, possibly, that it does not account for the *differences* in information-reduction strength of the various H relations. That is, some H relations reduce certain formulas more than do others. These differences are expressed in the redundancy formula below, where the numerator is in effect identical to a summation of the H relations weighted by their reduction strengths.

Redundancy (R′)

Redundancy is defined by

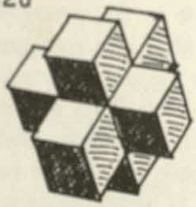
$$R' = \frac{I\,\max - I}{I\,\max},$$

where I max is the number of raw data elements and I the information. For example, for the sequence 0, 1, 2, 3, 4, 4, 3, 2 , 1, 0, represented by $R\,\{(4\cdot(1))\int\}$, I max $= 10$ and $I = 4$, so $R' = 6/10$. However, this measure is rather unstable with repetition. Suppose the given sequence is repeated five times. Then $R' = 9/10$. For all visual patterns, this measure is around the value 1; it is therefore not discriminative, in contrast to the hierarchy measure.

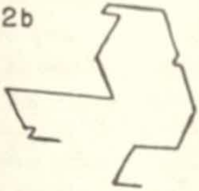
Substructure (G)

A representation of a pattern can be divided into groups of transformations, in such a way that between these groups H relations exist. A group may also contain H relations within it. These groups of transformations correspond to substructures of the pattern. Two points ought to be noted, for they will be important when we come to evaluate the perceptual phenomena discussed in the next main section. First of all, the structure of the formula, and therefore the nature of the subpatterns, is determined by efficiency principles and not by the subpatterns discernible in the pattern. Second, the number of subpatterns possible in the most efficient representation of a pattern is extremely limited in comparison to the number of all possible divisions of the pattern into subpatterns. For example, figure 12a contains subpattern 12b. But subpattern 12b is an arbitrary one and is therefore not identical to the *hierarchical* subpatterns of figure 12a, which can be found by looking in Table 1. There the representation of figure 12a (S-4) contains capital letters that refer to subpatterns represented in the same table; these latter are the hierarchical subpatterns of the figure.

12a



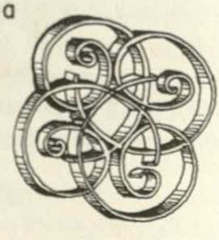
12b

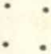


Hierarchical order (O)

Another aspect that must be mentioned is the hierarchical *order* of the subpatterns. For example, figure 13a has subpatterns 13b, 13c, and 13d. These three correspond to the three formulas (from top to bottom) that are given for the same main pattern as it is shown in Table 1. That the hierarchy between substructures is not arbitrarily reversible can be seen by comparing figures P and Q of Table 1, where the order of hierarchy is mirrored in the coded representation in the way that the superstructures are placed in the upper left of the formulas. This is because the superstructures contain the parameters of the substructures and need, therefore, to be solved first in reconstructing the figures from the formulas.

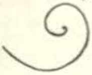
13a

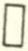




13b

13c





13d

PROPERTIES: EXPERIMENTAL AND PHENOMENAL EVIDENCE

The experimental applications are described in this section in about the same order as the description of the properties of the coding system: information, hierarchy, substructures, and then the hierarchical order of substructures. These applications subsume the remaining properties.

Information (I)

The most straightforward and obvious way to test the psychological applicability of the information measure is to ask subjects to judge the complexity of patterns. This was done by R. Young in the laboratory. Twelve students judged 16 three-dimensional patterns (see figures 14a–14p). Young found a correlation of $r > .94$ between the theoretical information (I) and the subjects' complexity judgments.

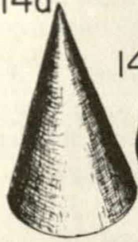
In another experiment, done to find out how much information is lost in the memory as a function of the information presented, 15 subjects were asked to copy eight three-dimensional pictures. The information content ranged from $I = 8$ to $I = 22$, with intervals of 2 units. In these patterns, 50% of the transformations were hierarchical, for reasons that will later become apparent. Each subject was shown four patterns for 30 sec each and then was required to draw the pictures from memory. The drawings were compared with the original patterns and the numbers of difference transformations (d) were calculated. The average d, or error, correlated with I at $r = .84$.

Not only the absolute amount but also the proportion of error was affected by I: the proportion of error, d, increased gradually from 0% to 30% as I went from 8 to 22. However, there was a positive correlation, $r = .87$, between *transmitted* information ($I - d$) and I itself. So there is no evidence for a maximum information capacity. Rather, the amount of information transmitted depends on the amount of stimulus information, at least for the stimulus values studied here.

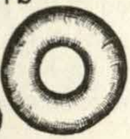
Also, a very complicated pattern ($I \sim 250$) was displayed to three groups of 10 subjects each for 30, 60, and 120 sec. The average transmitted information ($I - d$) here was 30, 60, and 106 respectively. This pattern's high level of hierarchy was, as will become apparent, very important to memory.

Finally, the phenomenal evidence suggests that the dimensionality of figures follows from the efficiency principles underlying the information measure. Figure 15a is interpreted as a two-dimensional figure rather

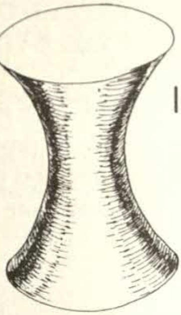
14a



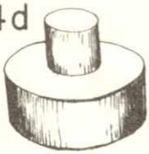
14b



14c



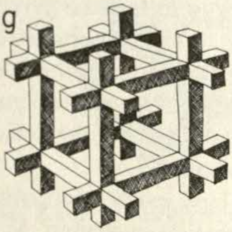
14d



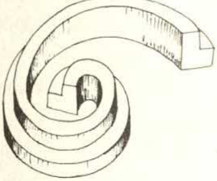
14e

14f

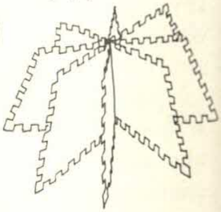
14g



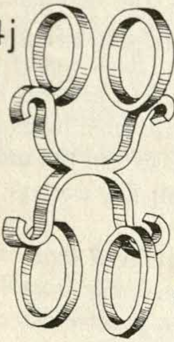
14h



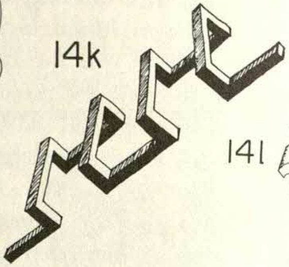
14i



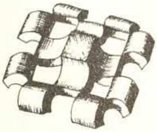
14j



14k

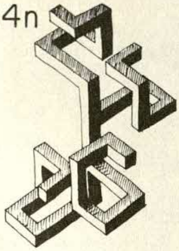


14l

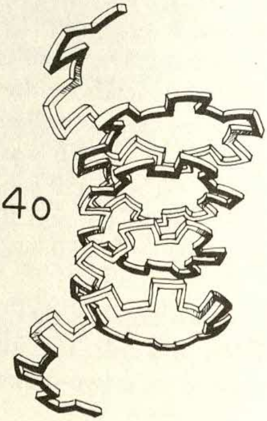


14m

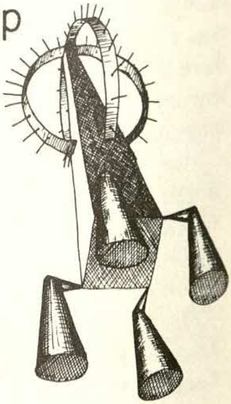
14n



14o

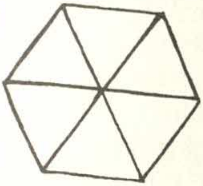


14p

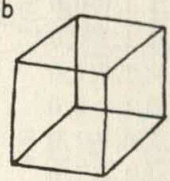


than as a three-dimensional one (Hochberg and McAllister, 1953). This can be explained by the different amounts of information of the different interpretations: as a two-dimensional figure, it contains but two information units (see figure R, Table 1); as a three-dimensional figure, it contains four information units. The reverse is the case with figure 15b: as a two-dimensional figure, it contains nine information units; as a three-dimensional figure, it contains only four units of information (see figure S, Table 1). So the efficiency principles also explain the phenomenal preference for the three-dimensional interpretation of figure 15b.

15a



15b



Auditory information (I)

This discussion starts with the binary auditory patterns (Table 3) constructed by Royer and Garner (1966). The response-defined measure of complexity they developed consists of the uncertainty of the onset of the response patterns. Vitz (1968) developed the *H*(run-span) as a measure of complexity of sequential patterns. The *H*(run-span) is the uncertainty of the runs in one period of a pattern. This measure correlated .92 with the onset certainty for the patterns in Table 3. The present measure of complexity (*I*) correlated .82 with the response uncertainty for the same patterns. Thus whereas the proposed coding system accounts for more complex patterns in which hierarchy is involved, the patterns in Table 3 were probably so simple that the overriding consideration here was metrical information, to which response uncertainty and *H*(run-span) are both applicable. It is still largely within the capacity boundaries of the listener to code the runs of the patterns as units, without further accounting for the relationships between them. Indeed, this type of coding is assumed in the Vitz model (1968).

Table 4 contains more complicated auditory patterns. The numbers at the left of Table 4 correspond to tones between which the intervals are perceptually indentical. Predictability judgments were obtained on these 'tunes' from 24 subjects, who were asked to listen and to understand the musical patterns in such a way that they could extrapolate

Table 3. Simple auditory patterns, from Royer and Garner, 1966

Pattern (one period)	Formula	Information
1, 0, 1, 0, 1, 0, 1, 0	}1,0{	2
1, 1, 0, 0, 1, 1, 0, 0	}2·(1,0){	3
1, 1, 1, 0, 1, 1, 1, 0	}{3·(1)}(0){	3
1, 1, 1, 1, 1, 1, 1, 0	}{7·(1)}(0)}	3
1, 1, 1, 1, 0, 0, 0, 0	}4·(1,0)}	3
1, 1, 1, 1, 1, 0, 0, 0	}5·(1),3·(0)}	4
1, 1, 1, 1, 1, 1, 0, 0	}6·(1),2·(0)}	4
1, 1, 1, 0, 1, 0, 1, 0	{(3,1,1);((1)(0))}	5
1, 1, 1, 0, 1, 0, 0, 0	{(3,1,0,0);((1)(0))}	6
1, 1, 1, 1, 0, 0, 1, 0	{(4,0,1);((1)(0))}	5
1, 1, 1, 1, 1, 0, 1, 0	{(5,1);((1)(0)){	4
1, 1, 1, 1, 0, 1, 0, 0	{(4,1,0);((1)(0))}	5
1, 1, 1, 1, 0, 1, 1, 0	{(4,2);((1)(0)){	4
1, 1, 1, 0, 1, 1, 0, 0	{(3,2,0);((1)(0)){	5
1, 1, 1, 0, 0, 1, 1, 0	{(3,0,2);((1)(0))}	5
1, 1, 1, 0, 0, 1, 0, 0	{(3,0,1,0);((1)(0)){	6
1, 1, 0, 1, 0, 1, 0, 0	{(2,1,1,0);((1)(0)){	6
1, 1, 0, 1, 0, 0, 1, 0	{(2,1,0,1);((1)(0))}	6
1, 1, 0, 1, 1, 0, 1, 0	{(2,2,1);((1)(0))}	5

the patterns. The time between the onset of the pattern and the moment at which the subject claimed he understood the pattern was measured. The correlation between these average 'understanding' times and the information (I) of the patterns was .86. The correlation between the understanding times and the Vitz H(run-span) measure was .62.

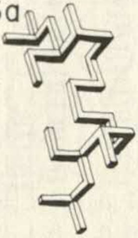
Unfortunately, the proposed coding language is without doubt inappropriate for music. This can be made clear by looking at the structures of some famous melodies from the viewpoint of the proposed language; no aesthetic principles emerge. Attempts of the type by Simon and Summer (1966) are more appropriate to the understanding of the sequential aspects of music. Their approach derives from musical theory, whereas the proposed coding language originated in the analysis of visual patterns and has so far been defied in its attempt to code music.

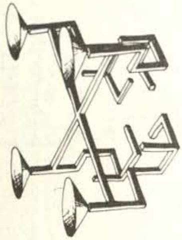
Table 1 Complex auditory patterns

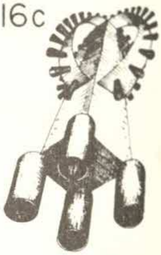
Pattern (one period)	$\sum^{24}$ 'understanding' time	Formula	Information
0, 1, 2, 3, 4, 4, 3, 2, 1, 0	277	}R(4·(1)∫)}	4
0, 0, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4	179	}3·(4·(1)∫)}	4
0, 4, 1, 4, 2, 4, 3, 4	195	}[3·(1)∫](4)}	4
0, 0, 1, 0, 2, 0, 3, 0, 4, 0, 4, 0, 3, 0, 2, 0, 1, 0	364	}(R{4·(1)∫})(0)}	5
0, 1, 1, 2, 3, 3, 4, 0, 0, 1, 2, 2, 3, 4, 4	302	}(1,2)(4·(1)∫)}	5
0, 2, 1, 3, 2, 4, 3, 3, 4, 2, 3, 1, 2	375	}(R{3·{2, −1}})∫}	5
2, 1, 0, 3, 2, 1, 4, 0, 2, 1, 3, 4	643	}(2)⁚({2,1}(0,3,4))}	6
0, 1, 2, 3, 4, 3, 2, 1, 2, 3, 2, 1	563	}((4,2)(3))(±)(1)} ∫	6
3, 2, 3, 2, 1, 2, 1, 0, 1, 0, 1, 0, 1, 2, 1, 2, 3, 2, 3, 4, 3, 4, 3	555	}∓{4·{1−1,1}}} ∫	6
2, 1, 0, 1, 0, 3, 1, 0, 4, 1, 0, 1, 0, 2, 1, 0, 3, 1, 0, 1, 0, 4, 1, 0	379	}(2,1)⁚((2,3,4){1,0})}	7
0, 4, 3, 4, 3, 1, 3, 4, 3, 4, 2, 4, 3, 4, 3, 0, 3, 4, 3, 4, 1, 4, 3, 4, 3, 2, 3, 4, 3, 4	420	}(0,1,2)[R{2·{4,3}}]}	7
4, 3, 4, 3, 4, 2, 3, 2, 3, 1, 2, 1, 2, 0, 1, 0, 1, 0, 2, 1, 2, 1, 3, 2, 3, 2	479	}∓[(4,3,3);(1)(2)]} ∫	7
0, 0, 1, 1, 2, 3, 2, 4, 4, 3, 3, 2, 4, 1, 0	566	}(2)⁚(4·(1)∫)[R{4·(1)∫}]}	8
0, 1, 2, 0, 2, 1, 0, 1, 0, 1, 2, 2, 2, 1, 0, 3, 0, 1, 2, 4, 2, 1, 0, 4, 0, 1, 2, 3, 2, 1, 0, 2, 0, 1, 2, 1, 2, 1, 0	1,192	}(R{0,1,2})[R{4·(1)∫}]}	8
0, 2, 1, 0, 1, 3, 2, 1, 2, 4, 3, 2, 4, 2, 3, 4, 3, 1, 2, 3, 2, 0, 1, 2	932	}±({2, −1, −1}(1,1,2))} ∫	8
2, 2, 1, 0, 1, 2, 2, 3, 3, 2, 1, 2, 3, 3, 4, 4, 3, 2, 3, 4, 4, 3, 3, 2, 1, 2, 3, 3	1,256	}{0,2·(−1,1),0}[2·(1, −1)]}∫)	9
0, 0, 1, 0, 0, 3, 1, 2, 1, 1, 0, 2, 3, 2, 2, 3, 3, 4, 3, 3, 0, 2, 2, 2, 2, 3, 1, 2, 1, 1	1,195	(0,3)(}{1, −1,0}(3·(1, −1))}∫)	9
0, 0, 1, 2, 1, 1, 4, 3, 3, 2, 3, 0, 0, 1, 2, 2, 1, 4, 3, 3, 2, 3, 3, 0, 1, 2, 2, 1, 4, 4, 3, 2, 3, 3, 0, 1, 1, 2, 1, 4, 4, 3, 2, 2, 3, 0, 1, 1, 2, 1, 1, 4, 3, 2, 2, 3	882	}(3,2)⁚((0)[±{1,1, −1,3}])} ∫	9

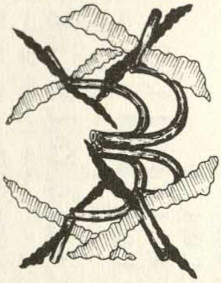
Hierarchy (H)

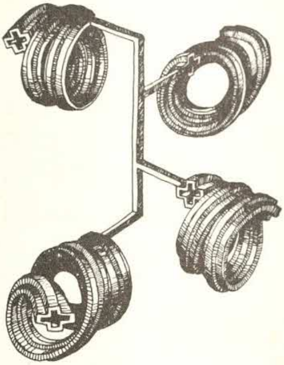
The relation between hierarchy and memory was examined using eight patterns varying in hierarchy ($H = 3$ to 10, see figures 16a–16h). Information content was constant at $I = 17$. Each subject was shown four patterns for 30 sec each and was then required to draw the pic-

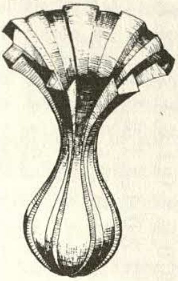

16a

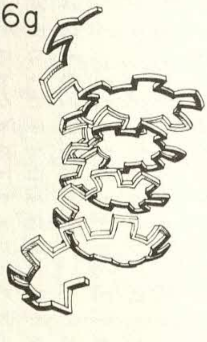

16b

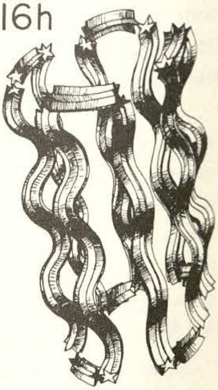

16c


16d


16e


16f


16g


16h

tures from memory. After this test, the drawings were compared with the originals and the numbers of difference transformations (*d*) were calculated. The average error, or *d*, correlated −.83 with *H*, and *H* thus correlated +.83 with remembered information.

This result does not follow from considerations up to this point. Rather, it is what is to be expected if the raw data (*I* max) are kept constant, for then the more the hierarchy, the less the information. Here, however, it was the information (*I*) that was kept constant, and hierarchy still showed a considerably favorable effect on memory. With *I* constant, this effect could not have been due to superior processing of redundancies, for the *I* measure itself excludes these repetitions. Something else was happening. It should be clear now why the amount of hierarchy was kept constant in some of the other experiments.

In another experiment, descriptions of the eight patterns and verbal judgments of their complexity were collected from the same 15 subjects. There was no clear correlation between the length of the subjects' descriptions and *H*. However, the correlation between *I* and the length of the descriptions was .86. There was no clear correlation between judged complexity and *H*. So *H* and *I* are clearly independent aspects.

From all these results it seems reasonable to hypothesize that more spontaneous attention is directed to the more hierarchical transformations in a pattern. Since *H* and *I* both have powerful effects on memory, one might still suspect some common aspect, but the phenomena described in the section immediately below clearly show that *H* and *I* as here defined are independent.

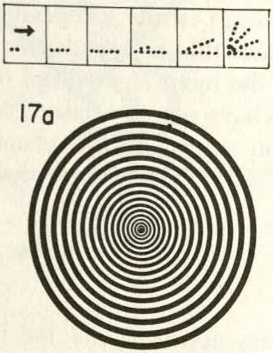
Substructures (G)

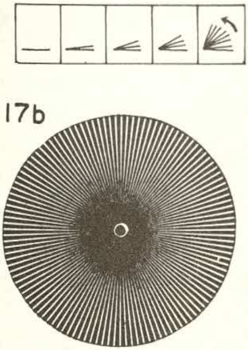
The point that is now discussed may appear trivial, but it gives the best evidence for the analogy between the coding language and the human coding system. The patterns theoretically derived from the efficiency principles are exactly those subpatterns that are spontaneously distinguished in perceiving the patterns without any special context. This means that the subpatterns easily detectable in the patterns (i.e., the perceptual constituents of the patterns) correspond to the theoretical subpatterns. These theoretical subpatterns are the groups of transformations shown at those entries in Table 1 where the groups are replaced by capital letters that refer to previous, more simple patterns; one can easily verify, by inspecting the table, that these simpler figures are the perceptual components of the more complicated patterns. The subjects' descriptions of the patterns in figure 16 also clearly contained these substructures as psychological units, and their judgments of perceptual similarity and their classifications of figures were mainly determined by the presence of common subpatterns.

Given that three-dimensional patterns are perceptually interpreted as hierarchical combinations of simple two-dimensional patterns, then the experimental evidence for two-dimensional patterns (Leeuwenberg, 1968) implies the experimental evidence for three-dimensional patterns.

Hierarchical order (O)

Twenty figures, together with their constituent substructures (for an example, see figure 13), were given to 20 subjects. The last substructures were given in a random order. Nearly all subjects were able to indicate which structures were functioning as superstructures and which as dependent structures, except for figures that consisted of patterns on a curved surface. These last are indeed ambiguous as far as hierarchy is concerned.


17a


17b

The main component of the aftereffect of figures 17a and 17b (from Mackay) can be explained by the coding procedure. The coding of 17b implies a rotating sequence of observations (see the sketch above the pattern). Likewise the coding of 17a implies not only a rotating but also a centrifugal sequence of observations (see the sketch above this pattern). The movements, which are perpendicular to the lines in the figures, are visible in the aftereffects of the figures. There is another way in which pattern 17b can be coded by a rotating sequence of observations, but this way would produce seven information units instead of the more efficient four.

A THEORY FOR THE CODING

We now attempt to bring together some a priori arguments for the correctness of choice of the proposed coding rules. This attempt will, however, be of a very noncommittal, speculative nature.

We shall start this way. It seems reasonable to assume that the visual system predicts what will be touched when a perceiver approaches an object. The visual system also warns the perceiver against possible injury. Therefore, knowledge of the forces in the environment is of vital importance. Now, the system that is most appropriate to assimilate these informative forces is the system that expresses the incoming information in terms of *deformative* forces—the sensorimotor system. Indeed, we might say that the possible 'deformation' of a perceiver by an object can be directly derived from the deformations in the sensorimotor system that scans the shape of that object. So we assume further, on this admittedly a priori basis, that the coding rules of visual perception are based on the coding rules of sensorimotor perception, and not the reverse. This notion is also in accordance with Pavlov's interpretation of the development of the brain, and therefore the visual system, as a secondary signal system, conditioned by a 'lower' information-processing system.

If our assumptions are correct, the visual 'analysis of shape' has to be expressed in the type of information in which the sensorimotor system codes the shape of an object, namely, in terms of the sequence of deformations that occur in the scanning system as it moves along the shape of an object. In other words, the structure of objects must be expressed by the visual system as movement and deformation.

However, this expression must also be efficient. This means that only the features that contain no repetition (are not redundant, and are therefore not predictable) may be used. This in turn means that the coding rules themselves must be efficient. The greater the number of features that can be predicted, then the smaller the number of rules required to predict them; and the less the rules are themselves predictable from other rules, the more general the rules. So the most general rules are the most efficient ones for the coding system. These rules have to apply to the objects to which the visual system is directed. In other words, they must be based on the natural laws that play a part in the structure of objects. On the other hand, these laws describe the course of movement, as we have already noted. This implies that the motion laws that play a part in the structure of objects must answer to the coding rules. Let us proceed on that basis.

The question now is how we are to conceive of motion laws being involved in a form that is not of a dynamic nature but, rather, preeminently static. The only way in which this can be understood is to regard the form of an object purely and solely as the resultant of a dynamic *forming* process. This physical forming process of an object may be imagined as a step-by-step sequence of more or less integral, dynamic interactions with other objects, in such a way that every deformative movement has an effect on the result of the previous deformations. Each of these dynamic interactions may be described by the motion laws. Consequently, we represent the form of an object as the succession of motion laws that have been involved, consecutively, in the growth of that form. As a conclusion, we state that visual encoding is the most probable reconstruction of the physical history of an object, whereby the information units correspond to the series of motion laws entailed.

If this conclusion is correct, one would expect the informational transformations (n, $\int$, R) to be in accordance with Newton's laws of motion. As regards their number, three, there is indeed agreement. (The ± transformation can be reduced to the R transformation.) As regards the content of the laws and the transformations, there is, in our opinion, also a correspondence. We now attempt to make this correspondence plausible, although we admit that tolerance and perhaps some fantasy on the part of the reader would be highly appreciated.

The first law of motion, in brief, reads: Given a body in motion, it continues with the same motion until a counterforce acts on it. In a remote sense, this could be expounded as: If a certain process (e.g., the printing of copies) has been initiated, then its continuation, or iteration, is expected. A stop to this process is unexpected and therefore informative. Thus the limitation of a given datum corresponds to one information unit. And when a number limits a certain dimension, the number corresponds to an information unit.

The second law of motion reads, in brief: The exercise of a certain force on an object is reflected in a constantly renewed addition of a certain velocity to the object's given velocity in the preceding moment. In other words, the invariance of force is expressed only if differences appear between velocities at each pair of consecutive moments of time. The integration transformation, $\int$, discloses these differences. The use of this transformation might then, considering its use in this natural law, be regarded as justified. Since this transformation is but one of the transformations that might be chosen, since it therefore has alternatives, it must correspond to an information unit.

The third law of motion states: Action is equal to reaction. Here reaction presupposes a force oppositely directed to the force of the action. To what degree this law is responsible in the construction of objects for the symmetry appearing in the final shape of objects is not quite clear. In any case, the law itself contains a symmetry that may justify the use of symmetry as a coding rule. The information unit should not be allotted to the structure that occurs in symmetrical form, but to the fact that there is a certain instance of symmetry.

The basic assumption is, again, that the perceptual coding language owes its rules to physical laws of the objects. The visual system refers to the shapes of the objects in the environment of the perceiver. To what aspect of the environment does the auditory language refer when it codes sequences of tones? For an answer to this question, the meaning of tone sequences in the environment of the perceiver has to be more closely analyzed. Now, tone sequences are usually produced by other living systems and not by inorganic processes. This makes it probable that sounds are not appropriate for the neutral description of situations but rather for the expression of those situations insofar as they are relevant to living systems. Hence, sounds form the language for the way in which outside situations are received and interpreted by living systems.

In other words, sounds form the language of feelings, emotions, and their derivations. Clearly, as the world of neutral objects is different from the world of feelings and emotions, the proposed coding language is not entirely applicable to auditory patterns. That it applies to a certain extent is because there is a certain overlap between the neutral objects and their interpretations for a living system. In any case, in order to make an appropriate coding language for sounds, it would be necessary first to analyze and categorize feelings, emotions, and their underlying drives.

Finally, since the number of *visual* patterns that can be stored in long term memory is so extremely large (Shepard, 1967), the coding for these visual objects is based on efficiency principles. Consequently, the coding system is applicable only to long-term storage. In contrast to the long-term capacity stands the very limited amount of patterns that can be assimilated at the same time or evoked in the imagination. So the imaginal representation is, in content capacity, quite different from the long-term memory. Hence the imaginal representation must be different from the long-term coding (Moran, 1969).

Suppose, in order to clarify this last point, that a subject is asked to specify the shape of an intersection of an ellipsoid and a certain flat surface. He cannot solve the problem before making an image of the two shapes. Analogously, it is impossible to derive directly, in a mathematical way, the solution to this problem given only the formulas of the two shapes. Only when the ellipsoid and the surface are constructed from the formulas is it possible to see what their intersection looks like. From this analogy it seems reasonable to assume that the reconstruction of patterns from their long-term formulas occurs in image space. This image space must be analogous to the real three-dimensional space. In other words, the elementary 'grains' of the image space must correspond, one-to-one, to the points in the space in which real shapes exist. So the image contains much less information but (Attneave, 1964) represents the information in a much more developed, demonstrative, and therefore more redundant way than the long-term memory.

Notes

This research was supported by the Niels Stensen Foundation and the Netherlands Foundation for Scientific Research (Z.W.O.). Received for publication September 8, 1970.

1. The direction of the outerproduct is orthogonal to the surface constituted by the two last angles, in this case, 14 and 67 deg. However, there are two possibilities now. To know which is the intended one, it is necessary to place a screwdriver orthogonal to the mentioned surface and to turn it in accordance to the 14, 67 deg rotational orientation. The direction in which the screwdriver moves, when it is turning around as described, is the direction of the outerproduct.
2. Although the interaction between metrical information (M) and structural information (I) does not disturb the structural-information measures, it is still safer to reduce M as much as possible. The experiments discussed below were, we hope, liberated from M by keeping low integral values on every dimension. Also, information was derived from subjects' drawings by accounting only for the rank values of lengths and angles and ignoring absolute variations.
3. The arguments of this function are on the left-hand side. For example, $(a,b,c,) \sin \int = (\sin a, \sin b, \sin c) \int = 0, \sin a, \sin a + \sin b, \sin a + \sin b + \sin c$.

References

Attneave, F. 1954. Some informational aspects of visual perception. *Psychological Review* 61:183–193.

Attneave, F. 1967. Criteria for a tenable theory of form perception. In *Models for the perception of speech and visual form*, ed. W. Wathen-Dunn, pp. 56–67. Cambridge, Mass.: The M.I.T. Press.

Cattell, R. B. 1968. Are I.Q. tests intelligent? Psychology Today, March, p. 56.

Handel, S., and Garner, W. R. 1966. The structure of visual pattern associates and pattern goodness. *Perception and Psychophysics* 1:33–38.

Hochberg, J., and McAllister, E. 1953. A quantitative approach to figural "goodness." *Journal of Experimental Psychology* 46:361–364.

Leeuwenberg, E. 1968. *Structural information of visual patterns*. The Hague: Mouton.

Leeuwenberg, E. 1969. Quantitative specification of information in sequential patterns. *Psychological Review* 76:216–220.

Mackay, D. M. 1950. Quantal aspects of scientific information. *Philosophical Magazine* 41:289–301.

Moran, T. 1968. The grammar of visual imagery. Unpublished manuscript, Carnegie-Mellon University.

Pavlov, I. P. 1960. *Conditioned reflexes*. New York: Dover.

Pollack, I., and Klemmer, E. T. 1954. The assimilation of visual information from linear dot patterns. Air Force Cambridge Research Center technical report (July), pp. 54–66.

Restle, F. 1971. Theory of serial pattern learning: Structural trees. *Psychological Review* 77(6):481–495.

Royer, F. L., and Garner, W. R. 1966. Response uncertainty and perceptual difficulty of auditory temporal patterns. *Perception and Psychophysics* 1:41–44.

Shannon, C. E., and Weaver, W. 1949. *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shaw, A. C. 1968. The formal description and parsing of pictures. A doctoral dissertation that appeared as Stanford University Computer Science Department SLAC report 84.

Shepard, R. N. 1967. Recognition memory of words, sentences, and pictures. *Verbal Learning Behavior* 6:156–163.

Simon, H. A. 1969. *The sciences of the artificial*. Cambridge, Mass.: The M.I.T. Press.

Simon, H. A., and Summer, R. K. 1968. Pattern in music. In *Human judgment*, ed. B. Kleinmutz, ch. 8. New York: Wiley.

Vitz, P. C. 1968. Information, run structure, and binary pattern complexity. *Perception and Psychophysics* 3:275–280.

continued from page 460

Hutt, S. J., and Hutt, Corinne, eds. *Behaviour Studies in Psychiatry*. New York: Pergamon Press, 1970. Pp. ix, 221. $6.75.

Johnson, Marjorie Seddon, and Kress, Roy A., eds. *Thought-Language-Reading*. Proceedings of the 1969 Annual Reading Institute at Temple University, vol. 3. Philadelphia: Temple University Reading Clinic, 1970. Pp. 108. $4.50.

Kakar, Sudhir. *Frederick Taylor: A Study in Personality and Innovation*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. xi, 221. $6.95.

Kanfer, Frederick H., and Phillips, Jeanne S. *Learning Foundations of Behavior Therapy*. New York: Wiley, 1970. Pp. xi, 635. $10.95.

Kirk, G. S. *Myth: Its Meaning and Functions in Ancient and Other Cultures*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 299. $7.95.

Kragh, Ulf, and Smith, Gudmund, eds. *Percept-Genetic Analysis*. Stockholm: Gleerups Lund, 1970. Pp. 405.

Kramer, Ernest. *A Beginning Manual for Psychotherapists*. New York: Grune & Stratton, 1970. Pp. 132. $5.75.

Kubler-Ross, Elizabeth. *On Death and Dying*. New York: Macmillan, 1970. Pp. viii, 260. Paperback, $1.95.

Kugelmass, I. Newton. *The Autistic Child*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xii, 371. $16.50.

Laing, R. D., and Esterton, A. *Sanity, Madness, and the Family*. Baltimore: Penguin Books, 1970. Pp. 282.

Legge, David, ed. *Skills*. Baltimore: Penguin Books, 1970. Pp. 378. Paperback, $1.95.

Lehmann, Heinz E., and Ban, Thomas A. *Pharmacotherapy of Tension and Anxiety*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xii, 129. $7.50.

Leuner, H. *Katathymes Bilderleben*. Stuttgart: Georg Thieme, 1970. Pp. 76. DM 9.80.

Levin, Harry, and Williams, Joanna P., eds. *Basic Studies on Reading*. New York: Basic Books, 1970. Pp. x, 277. $10.00.

Libaw, F. B., ed. *1970–71 Mentorex Test Item Catalog for Introductory Psychology*. Beverly Hills, Calif.: Cognitive Systems, 1970. Pp. xiii, 734.

Libaw, F. B., and Kirschenbaum, Jack. *Mentorex for Introductory Psychology: Student Manual*. 2d ed. Beverly Hills, Calif.: Cognitive Systems, 1970. Pp. xiii, 213.

Limner, Roman Rechnitz. *Sex and the Unborn Child*. New York: Julian Press, 1969. Pp. 229. $6.95.

Mackay, R. Stuart. *Bio-Medical Telemetry*. 2d ed. New York: Wiley, 1970. Pp. 525. $14.95.

Mannoni, O. *Freud*. New York: Random House, Pantheon Books, 1970. Pp. 215. $6.95.

Marchais, P. *Glossaire de psychiatrie*. Paris: Masson, 1970. Pp. 238. Fr 40.

Marchais, P. *Psychiatrie et methodologie*. Paris: Masson, 1970. Pp. 212.

Maslow, Abraham H. *Motivation and Personality*. 2d ed. New York: Harper & Row, 1970. Pp. xxvii, 369. $5.95.

McCammon, R. W. *Human Growth and Development*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 295. $9.00.

continued on page 376

American Journal of Psychology
1971, Vol. 84, No. 3

RECOGNITION MEMORY FOR PHOTOGRAPHS OF FACES

Ruth Ellen Galper and Julian Hochberg
Yeshiva University and Columbia University

Recognition memory was compared for faces seen in positive, in negative, and in two different expressions. Pure pattern characteristics plus those properties that differ in smile and resting state cannot fully account for the accuracies obtained. But the characteristics that differ in smiling and nonsmiling expressions did play some role in recognition, which suggests that expressional variation offers a technique for investigating the properties by which faces are discriminated and remembered.

Hochberg and Galper (1967) demonstrated that although subjects can recognize still photographs of faces with surprising accuracy, recognition is significantly poorer for inverted faces, a finding corroborated recently by Yin (1969). Galper (1970) found that recognition accuracy was also impaired when photographs of faces were presented in negative. These results indicate that recognition accuracy for faces is better than for patterns of identical geometrical complexity (i.e., the inverted or negative photographs) and that this superiority rests at least in part on information that becomes inaccessible when a photograph is inverted or its brightness relations reversed.

More subtle (and more natural) transformations may also affect face recognition, to the extent that viewers rely on those characteristics that are altered by changes in expression. If viewers could not identify which of two different expressions of the same person was previously presented, this would be evidence that those changes in pattern associated with the difference in expression do *not* include the characteristics which mediate recognition. As a step toward determining the stimulus attributes in terms of which photographs of faces are perceived and remembered, therefore, the present experiment investigated the effect of an expressional variation (resting state versus smile) on recognition accuracy, and compared that effect with the accuracy obtained using previous procedures.

METHOD

—Subjects—One hundred and three (52 male, 51 female) undergraduate students at New York University served as subjects.

—Materials—Thirty-two photographs of females were selected at random from a college yearbook, after those with glasses or distinctive clothing or backdrop had been excluded. These pictures were then rephotographed and prepared as 35-mm slides. Half the photographs, chosen at random, were prepared in positive; half were prepared in negative. In addition, each of eight female models, ranging in age from 21 to 25 years, was photographed twice, once with a relaxed, unsmiling expression, and once with a full smile. (All photographs were full-face.) These 16 photographs were converted (in positive) to 35-mm slides. The head-to-field ratio and the brightness and contrast properties of the photographs of the models were carefully matched to those of the yearbook photographs, so that there was no discernible difference between the slides from the two sources.

—Procedure—The subjects, tested in groups of 30 to 35, were asked to view an inspection series of 24 slides with the aim of later being able to identify those pictures as having appeared in the series. Five days later, they viewed a test series of 24 slides, each of which showed a picture from the inspection series paired side-by-side with a picture not previously seen. The subjects indicated, for each test-series slide, which picture they recognized from the inspection series.[1]

The expressions of the faces in the *inspection series* varied only along the continuum from 'neutral,' or resting state, to full smile, with approximately half of the faces representing these extremes. *Set A* presented eight faces in positive, with expressions varying from resting state to full smiles; *set B* presented eight faces in positive, with four faces in the resting state and four in full smiles; and *set C* presented eight faces in negative, with expressions varying from resting state to full smiles. The sequence of presentation was semirandomly determined, with the restriction that no more than two successive faces be from the same set. Each slide was exposed for 7 sec, with 3-sec intervals between slides.

There were three conditions in the *test series.* In the *positive two-person condition,* eight slides were presented in positive, and each slide consisted of a face from set A paired with another person not previously seen. In the *two-expression condition,* eight slides were presented in positive, and each slide consisted of a face from set B paired with another view of the same stimulus person bearing an expression (resting state or full smile) different from that which had appeared in the inspection series. In the *negative two-person condition,* eight slides were presented in negative, and each slide consisted of a face from set C paired with another person (in negative) not previously seen.

The sequence of presentation of test-series slides corresponded to that in which the previously seen faces had been viewed in the inspection series. Pairing was such that the correct answer was 'right' for half of the slides in each condition, and 'left' for half. Each slide was exposed for 7 sec, with

3-sec intervals between slides. It should be noted that the subjects were given no specific information, prior to viewing the inspection series, on the nature of the recognition task. The word 'picture'—rather than 'person'—was consistently employed in the instructions, so that the subjects were neither systematically biased to expect either intra- or interperson comparisons, nor technically deceived by the instructions.

RESULTS

The mean number of errors in the positive two-person condition was .88; in the two-expression condition, 1.91; and in the negative two-person condition, 2.58 (see Table 1). These error scores were compared with the error score of 4.0 that would be expected on the basis of chance ($z = -28.3636$, -13.9333, and -10.1429, respectively; $p < .001$ in all cases). The results of the analysis of variance of the error scores, after square-root transformation, are given in Table 2.

Table 1. Means, medians, and standard deviations of error scores

Condition	Mean errors	Median errors	*SD*
Positive two-person	.88	.13	2.01
Two-expression	1.91	1.39	2.62
Negative two-person	2.58	2.00	2.88

Table 2. Summary of analysis of variance of transformed error scores

Source	*SS*	*df*	*MS*	*F*	*p*
Conditions	14.66	2	7.33	66.64	< .001
Subjects	27.51	102	.27	2.46	< .005
Residual	21.32	204	.11		
Total	63.49	308			

A Duncan new multiple-range test was applied to the differences between each of the three possible pairs of means. It yielded highly significant differences ($p < .005$) for each comparison.

DISCUSSION

As in earlier studies involving a variety of stimulus manipulations (Hochberg and Galper, 1967; Galper, 1970), recognition accuracy in the present investigation was significantly better than chance in all conditions. Recognition accuracy in the two-expression condition, however, was relatively poorer than that in the positive two-person condition,[2] and the negative two-person condition was poorest of all. Because recognition accuracy in the negative condition was significantly poorer than in the other conditions, we conclude that pure pattern recognition —that is, without reference to responses made to the faces as faces— cannot account for the degree of accuracy obtained in the other two conditions. The fact that recognition accuracy in the two-expression condition was lower than in the positive two-person condition, yet higher than in the negative condition, indicates that faces are not remembered and recognized solely by the combination of those characteristics which survive in negative (e.g., face shape, hair style) and those which are altered by the expressional variation used here.

The fact that accuracy in the two-expression condition was significantly better than chance (and better than that in the two-person negative condition) indicates that the expressional variation does alter aspects of the stimulus pattern which are used in remembering faces— that the characteristics which differ in resting state and smile play some role in face recognition. Further manipulation of expressional variation thus appears to be a promising means of determining the nature of the information on which face recognition is based.

Notes

This research was supported in part by National Science Foundation Grant GB 5270 to the second author. Received for publication September, 1970.

1. This procedure is adapted from that developed by Shepard (1967) for use with nonfacial stimuli.
2. A preliminary version of this experiment, using 23 subjects, an inspection series of 35 faces, and a test series of 16 faces (all in positive), yielded results entirely consistent with these findings. Recognition memory in the two-expression condition was significantly better than chance ($t = 9.50$, $df = 22$) and significantly lower than in the two-person condition ($t = 2.74$, $df = 22$).

References

Galper, R. E. 1970. Recognition of faces in photographic negative. *Psychonomic Science* 19:207–208.

Hochberg, J., and Galper, R. E. 1967. Recognition of faces: I. An exploratory study. *Psychonomic Science* 9:619–620.

Shepard, R. N. 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6:156–163.

Yin, R. K. 1969. Looking at upside-down faces. *Journal of Experimental Psychology* 81:141–145.

American Journal of Psychology
1971, Vol. 84, No. 3

TACHISTOSCOPIC HEMIFIELD EFFECTS WITH HEMIFIELD REPORT

Max Coltheart and Barbara Arthur
University of Waterloo, Canada

When subjects were postexposurally cued to report either the left or right half of a centrally fixated row of letters, and the exposure duration was kept short enough to preclude scanning by eye movements, there was no hemifield superiority in accuracy of performance. The finding is evidence against the view that briefly exposed, centrally fixated letters are processed from left to right.

Heron (1957) found that when a row of letters is displayed tachistoscopically with the subject fixating in the center of the row, more letters are correctly reported from the left of the fixation point than from the right. He also found that if all the letters in the display are on one side of the fixation point, performance is better when the letters are to the right of the fixation point than when they are to the left.

Heron proposed that both findings could be explained in terms of a property of the process involved in tachistoscopic recognition, namely, that it proceeds from left to right. With central fixation, the letters to the left of the fixation point will benefit because they will be processed first. With fixation at one end of the row, performance will be worse when fixation is to the right of the letters because the left-to-right scan cannot begin until after attention has skipped to the leftmost letter; if fixation is to the left of the letters, the scan can begin at once.

Coltheart and Merikle (1970) argued that the left-hemifield superiority obtained with central fixation could be due to the left-to-right order of report normally adopted by a subject (Bryden, 1960), rather than to the existence of a left-to-right scan. They showed that under conditions where order of report could not produce a bias favoring one of the two visual hemifields, the left-hemifield superiority no longer occurred. They concluded, therefore, that briefly displayed and centrally fixated rows of letters are not scanned from left to right and that the results previously used as evidence for the existence of such a left-to-right scan were due instead to the left-to-right report order used by the subjects.

This explanation of hemifield differences in terms of order of report does not apply directly to experiments in which letters are exposed in only one hemifield at a time. Order of report, however, still may be important in this situation. For example, if the subject attempts to report in a left-to-right order under these conditions, he will find it more difficult to do so when the letters are presented to the left hemifield, because here the letter he wishes to report first will have been presented rather far from the fovea and hence will be difficult to see. When the letters are presented to the right hemifield, the letter he wishes to report first will also be the letter closest to the fovea—the letter he sees best. It remains to be seen whether this explanation has any force; the situation will become clearer after the completion of research (now in progress) using partial-report techniques with letters presented to one visual hemifield.

One useful technique for investigating these issues is to present material in both hemifields but to require report of the material from only one hemifield, specifying which hemifield after stimulus presentation to prevent the subject's favoring the cued hemifield during presentation (by attending to it selectively or by changing his fixation). This technique was used by Fitzgerald and Marshall (1967). In their experiments, a row of eight letters was presented for 300 msec, with fixation in the center of the row. At the offset of the letter row, a tone was presented for 40 msec. A high tone meant that the subject should report only the right four letters; a low one, the left four letters. A full-report group was also run; for them, four letters were presented, two on each side of the fixation point, and each subject was asked to report as many of the four as possible.

Fitzgerald and Marshall found that their partial-report group reported the letters to the right of fixation more accurately than those to the left. They correctly point out that this is inconsistent with Heron's postulation of a left-to-right postexposural scan. If the left-hemifield superiority observed under full-report conditions with central fixation were due to the existence of an automatic left-to-right postexposural scan, a left-hemifield superiority should also have occurred under Fitzgerald and Marshall's partial-report condition, because by the time the scan reached the right-hand letters their representation in the fading visual trace would have partially or entirely decayed.

Fitzgerald and Marshall then proposed instead that the point from which the left-to-right scan begins was under the control of their postexposure cue, so that their subjects were only shifting attention to the left-hand end of the row when cued to report the left-hand four letters. When cued to report the right-hand four letters, the scan could begin immediately. They also suggested an alternative interpretation: that the scan always begins at the leftmost letter but has reached the right-hand letters at the time of the postexposure cue, and that it is this which favors the right-hand letters. This alternative seems scarcely intelligible. What is meant to be the status of the left-hand letters at the time of the cue? The fact that they have *already* been scanned should *favor* them, unless Fitzgerald and Marshall are suggesting that the subject has to go back and rescan them if the cue asks for the left-hand letters. Furthermore, if it takes 300 msec for the scan to reach the right-hand letters, one would not expect increases in exposure duration beyond 50 msec to have almost no effect on amount reported (Sperling, 1967). Also, why did they obtain a left-hemifield superiority with the full-report group if there is a scan from left to right and the most recently scanned items are best reported?

At any rate, both of the interpretations considered by Fitzgerald and Marshall depend upon the existence of a left-to-right postexposural scanning process. As is mentioned above, the existence of such a process has been challenged by Coltheart and Merikle (1970). If this challenge is to be taken seriously, an attempt must be made to explain Fitzgerald and Marshall's results in some other way. That is the purpose of our paper.

One unusual feature of their experiment was the very long exposure duration used, namely, 300 msec. Most experimenters use durations of 100 msec or less, their reason being a desire to prevent the subject's using eye movements to scan the display. The latency of a voluntary saccadic eye movement is between 120 and 200 msec (Dodge, quoted in Woodworth and Schlosberg, 1954; Ginsborg, 1953; Westheimer, 1954a, 1954b; Saslow, 1967) and the duration of a short saccade is less than 4 msec (Westheimer, 1954a). These figures indicate that an exposure duration of 300 msec is ample to permit the subject to take two looks at the stimulus, fixating at different points.

If the subject does move his eyes from the fixation point, which way will he move them? The fact that the eyes move from left to right in normal reading suggests that such a movement would be to the right, and Terrace (1959) did in fact report that any eye movements which occurred during or after the tachistoscopic exposure of a row of letters were predominantly movements to the right.

Thus, Fitzgerald and Marshall's subjects had time to move their eyes while the stimulus was present, so as to fixate it in two places; and what evidence that there is suggests that such a movement would usually be to the right. If this occurred, the four right-hand letters would benefit greatly from it, since after the eye movement the average distance of the right-hand four letters from the fovea would be reduced and the average distance of the left-hand four letters from the fovea would be increased. Thus the right-hemifield superiority observed by Fitzgerald and Marshall could have been produced by the occurrence of eye movements.

This last possibility could have been eliminated by substantially reducing the exposure duration. Fitzgerald and Marshall note that Winnick and Dornbush (1965) found no hemifield difference in accuracy of report when the subjects had postexposure instructions to report one-half of a centrally fixated row of letters. Winnick and Dornbush used an exposure duration of 150 msec, and their failure to find a hemifield difference is therefore consistent with the contention that the difference found by Fitzgerald and Marshall occurred because the exposure duration they used permitted eye movements to occur.

However, the form of postexposure cue used by Winnick and Dornbush was unsatisfactory, as Fitzgerald and Marshall point out. This cue was simply a spoken instruction and would therefore introduce an appreciable delay between the presentation of the stimulus and the presentation of the cue, a delay during which complete decay of a postexposural visual representation of the stimulus would have occurred.

In the experiment we report, the exposure duration was short enough to preclude eye movements during the presence of the stimulus, and the onset of the postexposure cue coincided with the offset of the letter display. In addition, under one condition, a visual-noise mask was presented at the offset of the letter display. If a postexposural left-to-right scan is responsible for the hemifield difference found by Fitzgerald and Marshall, this difference should be eliminated when a mask is used, since the mask abolishes the postexposural visual representation of the letter display, leaving the postexposural scan nothing to scan. This assumes that the mask does abolish the postexposural visual representation, an assumption which seems secure inasmuch as our criterion for judging the acceptability of a masking pattern was that the letters be unreadable when presented simultaneously with the mask.

METHOD

—Subjects—The subjects were 20 undergraduate students at the University of Waterloo, who served in the experiment to fulfill a course requirement.

—Material and apparatus—A set of 48 cards was prepared. Each contained a row of eight letters (Letraset 287 capitals), each letter subtending 13 min of arc vertically, with 28 min of arc between the centers of adjacent letters. The letters O and Q were not used. The remaining letters occupied each of the eight positions twice. No letter was used more than once on any card. The same random order of the 48 cards was used for all subjects.

The postexposure cue was delivered by means of a card containing two horizontal green bars, each 1 deg 38 min in length. The bars were located above and below the area occupied by the leftmost four letters (requesting a report of the left-hand four letters) or above and below the area occupied by the rightmost four letters (requesting a report of the right-hand four letters). Each cue was used on a randomly selected 24 of the 48 trials of the experiment. The same sequence of cue presentations was used for all subjects.

For subjects in the mask condition, the cue card also contained a horizontal strip of 'visual noise' (a 4 deg × 40 min random checkerboard of small black and white squares). This noise strip covered the entire area occupied by the letter row while it was displayed. When the visual-noise mask and a letter row were presented simultaneously, the letters in the row could not be read.

A Scientific Prototype GB three-field tachistoscope was used to present the letter rows and cues. The luminance of all three fields was 48 ftL, measured by a Pritchard Spectra photometer, when blank white cards were placed in the fields.

—Procedure—On each trial, after the experimenter's verbal ready signal, the subject was first presented for 2 sec with a fixation point (a small black square, located midway between the positions that would be occupied by the fourth and fifth letters of a row). A letter row was then displayed for 100 msec. Finally, a cue card, with or without a visual-noise mask, was displayed for 2 sec. A randomly selected 10 of the 20 subjects were presented with the mask on every trial; the cue cards for the remaining 10 were always plain white except for the cue bars.

All subjects were instructed to fixate the black dot, which would appear after the experimenter said 'Ready,' until the letters were presented. Then (for the no-mask group) when green bars appeared either to the left or to the right, or (for the mask group) when green bars accompanied by a black and white pattern appeared either to the left or to the right, the subject was to report verbally the four leftmost or four rightmost letters from the just-presented row, as indicated by the green bars.

RESULTS

Table 1 shows the means and standard deviations of the mean number of letters correct per trial as a function of visual hemifield and condition (mask versus no mask). Analysis of variance indicated that the mask impaired performance significantly [$F(1, 18) = 10.63, p < .005$]. Neither the hemifield effect nor the interaction approached significance, F being less than unity in both cases.

The absence of a hemifield difference in both the masked and unmasked conditions is evidence against the view that briefly exposed letter rows are processed from left to right, either during or after ex-

Table 1. Number of letters correctly reported (maximum possible, four) as a function of retinal locus and presence or absence of mask

		Visual hemifield	
		Left	Right
Mask condition	*M*	1.453	1.502
	SD	.275	.560
No-mask condition	*M*	2.024	2.209
	SD	.423	4.38

posure. The finding is consistent with the contention of Coltheart and Merikle (1970) that such hemifield differences as have been obtained in the past with centrally fixated letter rows were due to order of report. Since, in general, right-hand letters are reported after left-hand letters, the right-hand letters are subject to the effects of delay or output interference.

The existence of loss during output is directly implied by our results. If the postexposure cue were simply an instruction to report as many letters as possible, more letters would be reported from the left than from the right. In view of the absence of a hemifield difference in our experiment, however, this cannot be due to the fact that at the time of cue onset, more letters from the left than from the right are available. It follows from this that some of the right-hand letters available for report after stimulus offset become unavailable before they are reported. Therefore the hemifield differences found with full-report techniques are not a function of input (left-to-right processing) but of output (left-to right order of report).

The relationship between retinal locus and accuracy of report is shown in Figure 1. The function is roughly W-shaped. Such functions have previously been reported by Averbach and Coriell (1961), Haber and Standing (1969), and Coltheart and Merikle (1970), all of whom used partial-report techniques requiring the report of only one letter. The shape of these functions appears to be determined by two factors: visual acuity, which favors letters 4 and 5, and declines as distance from the fovea increases (Ludvigh, 1941); and spatial masking (interference produced by adjacent contours), which favors letters 1 and 8, since these are the only letters that do not have contours on both sides. (Haber and Standing, 1969, and Bouma, 1970, have both shown that the presence of adjacent contours impairs the detectability of a target letter.)

The curves obtained by Averbach and Coriell, Haber and Standing, and Coltheart and Merikle were symmetrical. The asymmetry of the curves shown in Figure 1 arises because a third factor is present here, one which is not present in studies requiring report of only one letter; that factor is order of report. Letters 4 and 5 are equivalent with respect to acuity and degree of spatial masking; but letter 5 is reported better than letter 4. This is doubtless because letter 5 will usually be given at the beginning. The same explanation can be given for the superiority of letter 1 over letter 8, for which the effects of acuity and spatial masking are also equivalent.

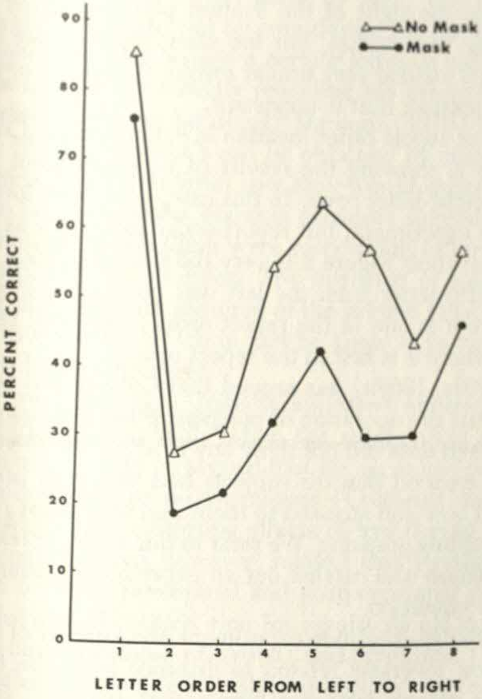


Fig. 1. Accuracy as a function of retinal locus

This analysis also applies to the results of Crovitz and Schiffman (1965), who presented eight-letter rows either to the left or to the right of the fixation point and plotted accuracy as a function of retinal locus. Their Figure 1 is very similar to our Figure 1. They concluded that relative location in the letter row was the most important determinant of accuracy.

In the experiment by Crovitz and Schiffman, however, as in most tachistoscopic experiments, location in the letter row was confounded with location in the report order. A byproduct of our experimental procedure is that these two variables are unconfounded. If location in the letter row were the important variable, our Figure 1 should have differed from theirs, since, for example, the item immediately to the right of the fixation point was the first in its row under their conditions and the fifth under ours. Despite this, performance with this item was similar in the two experiments. We suggest that the similarity is due to the fact that in both experiments this item is the first to be reported. In general, all items to the right of the fixation point occupied different row locations in the two studies, but the same report-order locations. Since the two studies yielded very similar results, it must be report-order location, not row location, that is important.

The importance of report-order location is confirmed by Crovitz and Schiffman's Figure 2, showing the results of an experiment involving central fixation of eight-letter rows. In this case, letter-row location was the same as in our experiment, but report-order location was different. The outcome is that their Figure 2 is very different from our Figure 1; for example, the fifth letter from the left was poorly reported in their experiment (where it is late in the report order) and well reported in our experiment (where it is first in the report order).

Hershenson (1969a, 1969b) has argued that Crovitz and Schiffman's data largely reflected the operation of postperceptual processes and that this is the reason their data did not show any relationship to visual anatomy. Hershenson ensured that his subjects had thorough prior knowledge of each letter row and stressed to them that they must report only what they saw, avoiding guessing. We refer to this as the 'foreknowledge condition.' Hershenson also carried out an experiment under the usual no-foreknowledge condition.

In the foreknowledge-condition experiment, the number of letters correctly reported was inversely related to distance from the fovea; that is, it followed the pattern expected from visual anatomy. Hershenson (1969b) took the view that this occurred because responses are isomorphic to the percept in the foreknowledge condition, whereas postperceptual processes play a large part in results obtained with the normal no-foreknowledge condition.

This last-mentioned view might have some plausibility if performance were worse under the no-foreknowledge condition, since the decrement could be attributed to forgetting during response. However, Figure 6 in Hershenson (1969b) shows that performance was much *better* in the no-foreknowledge condition. This is logically absurd on Hershenson's view, since it implies that items which were not represented in the percept (not reported in the foreknowledge condition) were often reported in the no-foreknowledge condition. The discrepancies are far too large to be ascribed to guessing: percent correct (estimated from Hershenson's Figure 6) for positions 1 and 2 were 67% and 65% for the no-foreknowledge condition, and 24% and 35% for the foreknowledge condition. Clearly, there was something bizarre about the task required of the subject in the foreknowledge condition, since he often denied seeing letters he could report perfectly well. The eccentric nature of this task can be brought out by considering the question, How does the subject *know* whether he saw a certain letter or not?

If the subject is not directly revealing his percept in the foreknowledge condition, what is he doing? Harcum (1970) has shown that results very like Hershenson's can be obtained by simply asking subjects which letters in a seven-letter row they would *expect* to see if the row were to be briefly exposed. Harcum therefore suggested that subjects in the foreknowledge condition were responding in terms of the demand characteristics of the situation. The intrusion of such an effect is not surprising, since the task required of the subject (that he report whether or not he saw a letter which in fact he knew to be there) must be very difficult for him to interpret.

Hershenson's technique initially appeared interesting because, with output interference unlikely when the subject had previously memorized all items, it could be argued that his accuracy/position curves, like those obtained with probe techniques (Coltheart and Merikle, 1970), would be free of the distorting effects of loss during report. The arguments presented above, however, suggest that Hershenson's attempt to distinguish between perceptual and postperceptual processes has been unsuccessful and, therefore, that his results should not be discussed in relation to ours and those of Crovitz and Schiffman (1965).

Our conclusion, then, is that plots of accuracy against retinal locus in tachistoscopic experiments reflect the joint influence of three factors: visual acuity (favoring letters close to the fovea), spatial masking (favoring end letters), and loss during report (favoring letters at the left of a row). We have given reasons for believing that neither position in the letter row as such (Crovitz and Schiffman, 1965) nor left-to-right processing (Heron, 1957; Fitzgerald and Marshall, 1967) is important in tachistoscopic performance, and our results appear to indicate that the hemifield difference reported by Fitzgerald and Marshall (1967) occurred because their subjects were able to move their eyes during the tachistoscopic exposure.

Notes

This research was supported by Grant APA-0350 from the National Research Council of Canada. Received for publication August 21, 1970.

References

Averbach, E., and Coriell, A. S. 1961. Short-term memory in vision. *Bell System Technical Journal* 40:309–328.

Bouma, H. 1970. Interaction effects in parafoveal letter recognition. *Nature* 226:177–178.

Bryden, M. P. 1960. Tachistoscopic recognition of nonalphabetic material. *Canadian Journal of Psychology* 14:78–86.

Coltheart, M., and Merikle, P. M. 1970. Are briefly processed letter rows processed from left to right? Presented at the meetings of the Canadian Psychological Association, Winnipeg.

Crovitz, H. F., and Schiffman, H. R. 1965. Visual field and the letter span. *Journal of Experimental Psychology* 70:218–223.

Fitzgerald, R. E., and Marshall, A. J. 1967. Left-right field differences with partial report of letters. *American Journal of Psychology* 80:370–376.

Ginsborg, B. L. 1953. Small voluntary movements of the eye. *British Journal of Ophthalmology* 37:746–754.

Haber, R. N., and Standing, L. 1969. Location of errors with a poststimulus indicator. *Psychonomic Science* 17:345–346.

Harcum, E. R. 1970. Perceptibility gradients for tachistoscopic patterns: Sensitivity or saliency? *Psychological Review* 77:332–337.

Heron, W. 1957. Perception as a function of retinal locus and attention. *American Journal of Psychology* 70:38–48.

Hershenson, M. 1969a. Perception of letter arrays as a function of absolute retinal locus. *Journal of Experimental Psychology* 80:201–202.

Hershenson, M. 1969b. Stimulus structure, cognitive structure, and the perception of letter arrays. *Journal of Experimental Psychology* 79:327–335.

Ludvigh, E. 1941. Extrafoveal visual acuity as measured with Snellen test letters. *American Journal of Ophthalmology* 24:303–310.

Saslow, M. 1967. Latency for saccadic eye movement. *Journal of the Optical Society of America* 57:1030–1033.

Sperling, G. 1967. Successive approximations to a model for short-term memory. *Acta Psychologica* 27:285–292.

Terrace, H. 1959. The effects of retinal locus and attention on the perception of words. *Journal of Experimental Psychology* 58:382–385.

Westheimer, G. 1954a. Mechanism of saccadic eye movements. *A.M.A. Archives of Ophthalmology* 52:710–724.

Westheimer, G. 1954b. Eye movement responses to a horizontally moving visual stimulus. *A.M.A. Archives of Ophthalmology* 52:932–941.

Winnick, W. A., and Dornbush, R. L. 1965. Pre- and post-exposure processes in tachistoscopic identification. *Perceptual and Motor Skills* 20:107–113.

Woodworth, R. S., and Schlosberg, H. L. 1954. *Experimental psychology*. New York: Holt, Rinehart and Winston.

American Journal of Psychology
1971, Vol. 84, No. 3

THE VISUAL PERSISTENCE OF A MOVING STROBOSCOPICALLY ILLUMINATED OBJECT

Robert Efron and David N. Lee
Veterans Administration Hospital, Martinez, California,
and University of Edinburgh, Scotland

A time-dependent visual illusion was studied to contrast the 'perceptual moment' hypothesis with the concept of 'persistence of vision.' The experimental results were formally consistent with either concept. However, the perceptual-moment hypothesis that fits the data is one in which the duration of the 'moment' is variable and determined by stimulus properties rather than by a constant, central, neural clock—a finding which undercuts the heuristic value of the hypothesis.

A small moving object viewed in the dark under low-frequency stroboscopic illumination gives rise to the illusion of a group of identical objects moving in tandem across the visual field. The number of objects in the group increases when the strobe frequency is increased. Since the objects in the group appear to be of similar contrast, the illusion cannot be explained in terms of afterimages (Karwoski and Crook, 1937; Allport, 1966).

Allport (1966, 1968) has attempted to explain this illusion in terms of a 'traveling moment' hypothesis. He distinguishes two versions of the perceptual-moment hypothesis. The first, which he calls a 'discrete moment' hypothesis, is exemplified by the model of Stroud (1955). In this view, perceptual information is time-sampled in single 'chunks' or 'frames' of approximately 100 msec, and these samples are consecutive, nonoverlapping, and discrete.

In the second version of the perceptual-moment hypothesis, exemplified by the model of William James (1890), the sampling epochs are not consecutive but are temporally overlapping. They are conceived as a 'moving' or 'traveling window' of approximately 100-msec duration. Allport describes this second version in terms of an analogy of an observer sitting in a moving train: his "field of view is always bounded by his own carriage window. New elements of the passing scene enter his view continuously from one side of the window, while others drop out of it at the other. Given a temporal rather than a spatial extension, the moving window analogy corresponds to the idea of a continuous 'travelling moment'" (1968, p. 396).

Allport has shown that the multiple-image phenomenon is inconsistent with the discrete-moment hypothesis but compatible with a traveling-moment hypothesis. Using the number of images perceived at any strobe frequency as a measure of the duration of the 'traveling moment,' his data reveal that the duration of the traveling moment is inversely related to the stimulus luminance. Mollon (1969) has shown that the number of images seen is also dependent on the spectral composition of the light. From these studies it appears that the perceptual moment—if indeed there is one—is of variable duration, is dependent on stimulus parameters, and may not be related to any fixed neurophysiological time constant, sampling epoch, or clock. This conclusion severely undercuts the heuristic value of the traveling-moment hypothesis, with its implicit reference to a central time-sampling mechanism.

There is, however, another reason for questioning the value of either version of the perceptual-moment hypothesis. Neither version deals explicitly with the phenomenon of persistence of vision. Because of the lack of explicitness, it is not clear if the phenomenon of persistence is subsumed by the perceptual-moment hypothesis or if the phenomenon of persistence is considered irrelevant to the theory. Expressed differently, has Allport measured the duration of the traveling perceptual moment or has he measured the duration of persistence, or are the two concepts to be considered identical? If they are identical, then what is the advantage of the concept of perceptual moments over the older, well-established concept of persistence?

It is the purpose of the present paper to show that the perceptual illusion of multiple images can be accounted for quantitatively by the phenomenon of persistence and without reference to any perceptual-moment hypothesis.

A recent series of studies of persistence of vision and audition has been performed using a method of intramodal (Efron, 1970a, 1970b) and crossmodal (Efron, 1970c) simultaneity judgments. By this technique, the magnitude of the perceptual onset and offset latencies could be compared. Persistence was defined as occurring when the perceptual offset latency was larger than the perceptual onset latency. It was found, for the particular stimuli and light-adaptation levels used, that the perceptual onset and offset latencies were equal when the stimuli were longer than 130 msec—that is, that these 'long' stimuli did not result in any persistence as defined above. However, for stimuli briefer than about 130 msec (the critical duration), the perceptual offset latency became larger than the onset latency and increased inversely with stimulus duration. It was also noted that the duration of the persistence for the 'brief' stimuli was inversely related to stimulus luminance.

From these results, it was anticipated that two or more consecutive brief stimuli (which have long persistence times) presented at different retinal loci would result in two or more perceptions that temporally overlap. That is to say, there would be some period of time during which the observer would be conscious simultaneously of more than one stimulus. This experimental paradigm is, in essence, identical with the conditions under which multiple images of a stroboscopically illuminated moving target are observed. It seemed appropriate, therefore, to compare the measure of persistence that might be obtained from an analysis of this illusion with the measure of persistence obtained in the previously reported experiments, which relied on an entirely different methodology.

METHOD

—Experimental paradigm—Consider a white disc on which there is a single black radial line. The disc is illuminated by a stroboscope and is rotated at a constant speed of about 1 rps, the conditions being such that the line is at nonoverlapping spatial positions at each successive strobe flash.

It is assumed that each strobe flash gives rise to a perception of the line at a particular locus at the time of the flash and that the duration of each perception is d sec. Since the duration of the strobe flash is of the order of 1 μsec, d can also be considered to be equal to the duration of the persistence. If the flashes are presented every $d/2$ sec, the perception produced by any one flash will temporally overlap (by 50%) the perception afforded by the next flash. This is illustrated in Figure 1. At any moment in time (see t_a, t_b, or t_c in Figure 1), the observer would be expected to perceive two lines, and these two lines should be seen to move in tandem in a circular path if the flashes continue to be presented at this frequency. Clearly, the same argument applies to the more general case where the interval between strobe flashes is d/n sec, n being a positive integer. In this case, a group of n lines should be seen moving together.

When n is not an integer, the situation is somewhat more complex. In this case, the observer would not necessarily be aware of the same number of lines at every moment. As an illustration, the case of $n = 10/3$ is represented in Figure 2. It may be seen that two-thirds of the time three perceptions temporally overlap (see t_a and t_c, Figure 2) and that one-third of the time *four* perceptions overlap (see t_b, Figure 2). If we make the reasonable assumption that the probabilities, in any one observational sample, of seeing three or four lines are 2/3 and 1/3 respectively, then it follows that with repeated observations, the expected *mean* number of perceived lines will be $2/3(3) + 1/3(4) = 3.3$.

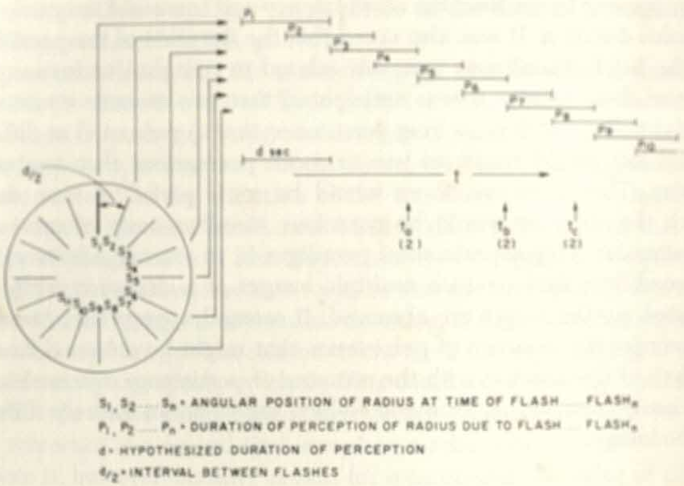

d/2
d sec
t
(2)
(2)
(2)
S1, S2 ... Sn = ANGULAR POSITION OF RADIUS AT TIME OF FLASH ... FLASHn
P1, P2 ... Pn = DURATION OF PERCEPTION OF RADIUS DUE TO FLASH ... FLASHn
d = HYPOTHESIZED DURATION OF PERCEPTION
d/2 = INTERVAL BETWEEN FLASHES

Fig. 1. Schematic representation of hypothesis, $n = d/2$

From such considerations we may deduce the following general prediction. If the interval between strobe flashes is d/k sec, where k is any number not less than unity and d (sec) is the persistence of the percept afforded by each strobe flash, then the expected mean number (n) of lines seen moving in a group will be equal to k. The corresponding frequency (f) of strobe flashes is k/d pps, and so the prediction may be simply written as the linear expression: $n = d(f)$.

This linear prediction was verified by the present experiments, and from the results the duration of perceptual persistence (d) could be computed. The duration of the persistence was measured for different luminance levels, it being expected, on the basis of other evidence (Allport, 1968; Efron, 1970), that the duration of perceptual persistence would increase as luminance decreased.

—Apparatus and procedure—A white disc of 27-cm diameter was mounted in a vertical plane on the shaft of an electric motor. On the disc was a single black radial line 2.5 mm wide. The disc was rotated at a constant speed of 78 rpm in the subject's frontoparallel plane, and was viewed at eye level from a distance of approximately 1 m. A General Radio 1539-A Stroboslave flash lamp, triggered by a Grass S8 stimulator, which had been accurately calibrated, was mounted behind and above the subject's head. Each flash was of 1.2-μsec duration and homogeneously illuminated the rotating disc. Using a method of brightness matching with a continuous source of illumination, the luminance of the disc was estimated to be approximately 1.2 ftL—a value close to that used in previous experiments (Efron, 1970c). The only other source of light in the matte-black experimental room was a small lamp used by the experimenter to record the subject's replies; this very dimly illumi-

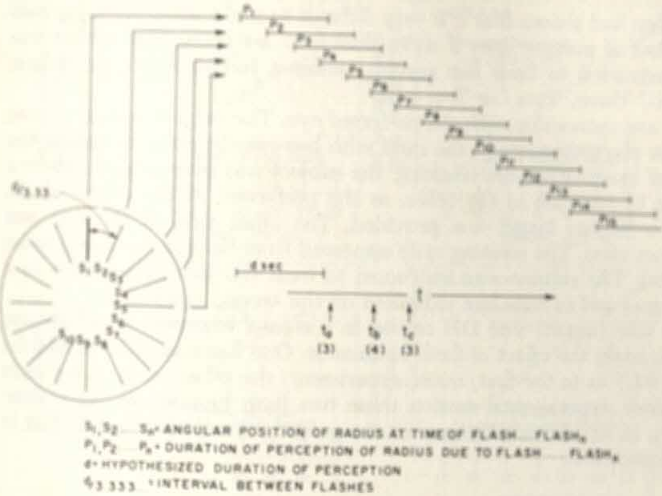

P1
P2
P3
Pn
d sec
t
tb
(3)
(4)
(3)
S1, S2 ... Sn = ANGULAR POSITION OF RADIUS AT TIME OF FLASH ... FLASHn
P1, P2 ... Pn = DURATION OF PERCEPTION OF RADIUS DUE TO FLASH ... FLASHn
d = HYPOTHESIZED DURATION OF PERCEPTION
1/3.333 = INTERVAL BETWEEN FLASHES

Fig. 2. Schematic representation of hypothesis, $n = 10/3$

nated the disc. In order to mask the sound of the strobe lamp, the subjects wore headphones into which white noise was delivered.

Five female employees of the hospital were used as paid subjects. They were all in the age range of 20–30 years. Three subjects were inexperienced with respect to any psychological experiments. Two had had previous experience as psychophysical subjects but had no knowledge of the theory or purpose underlying the present experiments. One of these latter, subject DH, had been used in previously reported experiments (Efron, 1970a, 1970b, 1970c).

At the start of the experiment the subject was shown the rotating disc under 25-pps stroboscopic illumination and was asked to report what she saw. Invariably, a sheaf of two or more radial lines was seen rotating as a group. The subject was then told that her task was to report, on each trial, the number of lines in the rotating sheaf. A short practice experiment, similar in form to the main experiment, was then run to familiarize the subject with the observational task.

Subjects KK, RB, DP, and GH served in the *main experiment*, which was divided into six sessions, each comprising 120 trials. In half the session the disc was rotated clockwise; in the other half, counterclockwise. Twelve strobe frequencies (10 to 32 pps in 2-pps steps) were used, chosen on the basis of pilot studies. Before each trial the strobe frequency was randomly preset at one of the 12 values, and the strobe was then turned on for only 1.5 sec. There were 10 trials at each frequency in a session. The random presentation of different strobe frequencies and the brief exposure to each frequency minimized any tendency the subject may have had to use an estimate of strobe frequencies as a cue to influence her judgment of the number of radii reported. Despite these precautions, it is possible that some cuing effect may still have influenced the results.

Pilot studies had shown that it is very difficult to make accurate judgments of the number of moving lines if more than four are seen. The subject was therefore instructed to limit her reports, under a forced-choice procedure, to 'one,' 'two,' 'three,' 'four,' or 'four-plus.'

Viewing was monocular with the preferred eye. The subject was instructed not to follow the movement of the radii with her eye. In order to reduce the possibility of inadvertent eye tracking, the subject was instructed to direct a steady gaze to one side or the other, as she preferred, of the center of the disc, but no fixation target was provided. The other side of the disc was shielded from view. The rotating radii appeared from behind the shield twice on each trial. The subject was instructed to treat the first appearance as a warning signal and to base her judgment on the second appearance.

Subjects GH (again) and DH served in a *second experiment*, which was designed to study the effect of flash luminance. One luminance level was the same (1.2 ftL) as in the first, main, experiment; the other was 1.5 log units lower. In each experimental session these two flash luminances were alternated every 10 or 12 trials. Otherwise the procedure was identical to that in the main experiment.

RESULTS

In the first experiment, where the luminance of the disc was approximately 1.2 ftL, there were four subjects. Since there was no significant difference between the results for counterclockwise and clockwise motion of the line, only the combined results are given in Figure 3. For these results, a report of four-plus was scored five.

It can be seen that the linear regressions of n on f for each subject strongly support the prediction that the mean number of radial lines reported (n) is proportional to the strobe frequency (f). The best estimate of the constant of proportionality is given by the regression slope. This value, a measure of the duration (d) of perceptual persistence, was approximately the same for all four subjects ($d = 133$, 137, 144, and 144 msec).

It was expected that the subjects might show more variability in their replies when presented with strobe frequencies which, according to the persistence hypothesis, should result in varying numbers of perceived radii. Although the selection of strobe frequencies was not ideally suited to this analysis, some evidence was obtained that the standard deviations did fluctuate as expected. The significance of this fluctuation was difficult to evaluate, since there was also present in every subject a more marked, systematic increase in the standard dēviation as the strobe frequency increased. This increase of standard deviation in the estimate of larger numbers of radii partially obscured the expected fluctuation of the standard deviation.

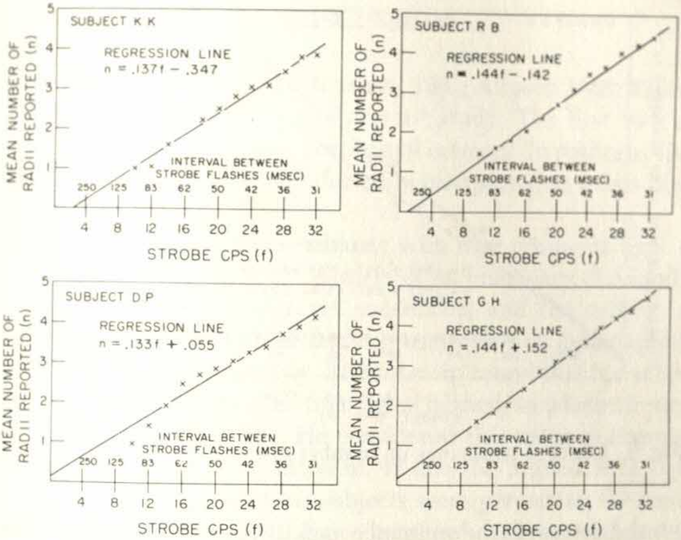


Fig. 3. Number of radii seen as function of strobe frequency

In the second experiment, where the luminance levels were 1.2 ftL and 1.5 log units lower (these two luminances alternating every 10 or 12 trials), there were two subjects. Again, there was no significant difference between the results for counterclockwise and clockwise motion, and so only the combined results are given in Figures 4 and 5.

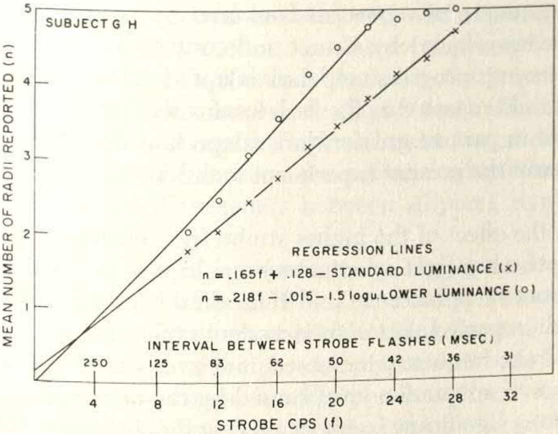


Fig. 4. Effect of luminance on number of radii seen; subject GH

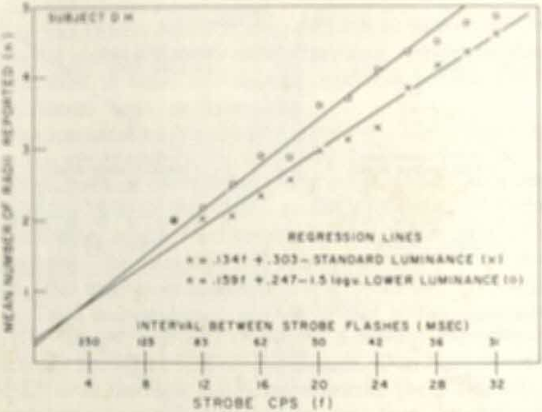


Fig. 5. Effect of luminance on number of radii seen; subject DH

Under the lower flash-luminance condition, the subjects reported more radial lines. The flattening of the lower luminance curves at the upper end is an artifact of scoring reports of four-plus as five. Accordingly, the upper two points were omitted in computing the regression equations given in the figures. The estimates of the duration (*d*) of perceptual persistence show an increase as flash luminance decreases. For subject GH, *d* increased from 165 to 218 msec (35 msec/log unit). For subject DH, *d* increased from 134 to 159 msec (17 msec/log unit).

It should be noted that the reduction in the luminance of the 1-μsec flash by 1.5 log units would also have altered the subject's level of light adaptation. Persistence of vision (as measured by critical flicker-fusion threshold) has been found by Granit and co-workers (1934, 1935) to increase as the eye is progressively dark adapted. Thus, the increase in the number of radii reported as the flash luminance was decreased could be due, at least in part, to greater dark adaptation. It was not possible with the design of the present experiment to differentiate between these factors.

In contrast, the effect of the higher strobe frequencies (which would tend to promote more light adaptation) would also be expected to reduce the amount of persistence and thus would tend to *decrease* the number of radii reported. Since an *increase* in the number of radii was found as the strobe frequency increased, it can reasonably be concluded that the change of adaptation level caused by the change in strobe frequency was not a significant factor in altering the persistence under the conditions of this experiment.

DISCUSSION

Three previous studies (Lichtenstein, 1961; Allport, 1968; Efron, 1970) are particularly pertinent to the present study. The first two of these were reported in terms of the 'perceptual moment' hypothesis. However, the results may be equally satisfactorily interpreted in terms of perceptual persistence.

Lichtenstein presented his subjects with four pin-point light sources spatially arranged to lie at the corners of an imaginary diamond. These four lights were flashed in rapid succession, and the subject was required to increase the rate of flashing until all four lights appeared to flash on and off simultaneously. Lichtenstein found that his subjects reported this end point when all four lights flashed, in whatever sequence, within a period of 125 msec. He considered this value to represent the duration of the perceptual moment. It may be argued with equal cogency that at this end point his subjects were perceiving the four lights at each and every instant in time, because of perceptual persistence. His value of 125 msec could thus be taken as an estimate of the duration of perceptual persistence under the luminance conditions he used (.05 ftL, 5 μsec).

Allport presented his subjects with 12 sequential horizontal traces on an oscilloscope screen, one just above the other. His subjects were to report when all 12 traces appeared to flash on and off simultaneously. He found that his subjects reached this end point when all the traces flashed within a period of approximately 90 msec and that the period was longer when the traces had a lower luminance. Allport attempted to relate this 90-msec period to the duration of a perceptual moment, but it too may be interpreted as a measure of perceptual persistence. That the value is lower than those obtained in the present experiments is likely due to the higher luminance Allport used. His lowest trace luminance was 16 ftL, whereas the highest luminance used in the present study was 1.2 ftL. Differences in the subjects' adaptation levels could also account for the discrepancy between Allport's results and those reported here.

In the previous study of persistence by the method of crossmodal simultaneity judgments (Efron, 1970c), a visual persistence time of 130 msec was obtained with stimuli of 3.4 ftL. Using the same subject (DH), the value of *d* obtained in the present experiments was 134 msec with stimuli of 1.2 ftL (see Figure 5). The previous experiments were performed under conditions in which the level of light adaptation was relatively constant. In the present experiments the light-adaptation level no doubt varied with each different strobe frequency. Despite this possible source of error, the close agreement in the estimate of visual persistence by two radically different methods gives support to the idea that the phenomenon of multiple images can be accounted for by persistence of vision and without recourse to any of the complex variants of the perceptual-moment hypothesis.

To account for these results in terms of perceptual moments, the duration of the moment would have to vary as a complex function of stimulus luminance, wavelength, state of adaptation, and so on. This modified concept of a perceptual moment stands in marked contrast to the original idea (implicit in *both* versions of the moment hypothesis) that the moment is a psychological consequence of a central information-sampling system operating on a fixed time base—a constant sampling epoch based on some biological rhythm or 'clock.' The modifications of the traveling-moment hypothesis that would make it formally consistent with the results of the present experiments would, at the same time, make it distinctly less interesting theoretically, since the duration of the alleged 'moment' would be primarily determined by stimulus parameters rather than by temporal parameters of any neurophysiological sampling mechanism.

Notes

The experiments were carried out when Dr. Lee was a research associate at Harvard Medical School. Support was provided by the Office of Naval Research under Contract Nonr 1866(52) with Harvard University and by a research grant from Boston City Hospital. Dr. Efron is in the Neurophysiology-Biophysics Research Laboratory of the Martinez facility. Received for publication April 14, 1970.

References

Allport, D. A. 1966. Studies in the psychological unit of duration. Unpublished doctoral dissertation, University of Cambridge.

Allport, D. A. 1968. Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology* 59:395–406.

Efron, R. 1970a. The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia* 8: 37–55.

Efron, R. 1970b. The minimum duration of a perception. *Neuropsychologia* 8:57–63.

Efron, R. 1970c. Effect of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics* 8:231–234.

Granit, R. 1935. Two types of retinae and their electrical responses to intermittent stimuli in light and dark adaptation. *Journal of Physiology* 85:421–438.

Granit, R., and Riddel, L. A. 1934. The electrical responses of light- and dark-adapted frogs' eyes to rhythmic and continuous stimuli. *Journal of Physiology* 81:1–28.

Granit, R., and Therman, P. O. 1935. Excitation and inhibition in the retina and in the optic nerve. *Journal of Physiology* 83:359–381.

James, W. 1890. *The principles of psychology*. New York: Holt.

Karwoski, T., and Crook, M. N. 1937. Studies in the peripheral retina: I. The Purkinje after-image. *Journal of General Psychology* 16:323–356.

Lichtenstein, M. 1961. Phenomenal simultaneity with irregular timing of components of the visual stimulus. *Perceptual and Motor Skills* 12:47–60.

Mollon, J. D. 1969. Two approaches to the perceptual moment hypothesis. Unpublished paper presented to Experimental Psychology Society, Oxford.

Stroud, J. M. 1955. The fine structure of psychological time. In *Information theory in psychology*, ed. H. Quastler, pp. 174–207. Glencoe, Ill.: Free Press.

continued from page 350

McCandless, Boyd R. *Adolescents: Behavior and Development*. Hinsdale, Ill.: Dryden Press, 1970. Pp. xii, 515.

McLean, Alan, ed. *Mental Health and Work Organizations*. Chicago: Rand McNally, 1970. Pp. 328. $5.75.

McNeil, Elton B. *Neuroses and Personality Disorders*. Englewood Cliffs, N.J.: Prentice-Hall, 1970. Pp. 176. $6.95; paperback, $2.75.

Medinnus, Gene R., and Johnson, Ronald C., eds. *Child and Adolescent Psychology: A Book of Readings*. New York: Wiley, 1970. Pp. ix, 662. $9.95; paperback, $5.95.

Mendel, Werner M., ed. *A Celebration of Laughter*. Los Angeles: Mara Books, 1970. Pp. 189. $4.95.

Meyer, V., and Chesser, Edward S. *Behaviour Therapy in Clinical Psychiatry*. New York: Penguin Books, 1970. Pp. 288. Paperback, $2.95.

Mill, John Stuart. *The Subjection of Women*. Cambridge, Mass.: The M.I.T. Press. 1970. Pp. xxix, 101. Paperback, $1.95.

Milner, Peter M. *Physiological Psychology*. New York: Holt, Rinehart and Winston, 1970. Pp. x, 531. $13.95.

Minium, Edward W. *Statistical Reasoning in Psychology and Education*. New York: Wiley, 1970. Pp. 465. $9.95.

Mink, Oscar G. *The Behavior Change Process*. New York: Harper & Row, 1970. Pp. 208. Paperback, $3.95.

Morrison, Denton E., and Henkel, Ramon E., eds. *The Significance Test Controversy*. Chicago: Aldine, 1970. Pp. xviii, 333. $12.50.

Moskos, Charles C., Jr. *The American Enlisted Man*. New York: Basic Books, 1970. Pp. 274. $7.95.

Mussen, P. H., ed. *Carmichael's "Manual of Child Psychology."* Vols. 1 and 2. 3d ed. New York: Wiley, 1970. $40.00 the set; Vol. 1, $30; Vol. 2, $20.

Neff, William D., ed. *Contributions to Sensory Physiology*. New York: Academic Press, 1970. Pp. xii, 220. $11.50.

Nordenstreng, Kaarle. *Toward Quantification of Meaning: An Evaluation of the Semantic Differential Technique*. Helsinki: Snellmanink, 1970. Pp. 35.

Norman, Donald A. *Memory and Attention: An Introduction to Human Information Processing*. New York: Wiley, 1969. Pp. ix, 201.

O'Connor, Elizabeth. *Our Many Selves*. New York: Harper & Row, 1971. Pp. xxi, 201. $4.95; paperback, $.95.

Paillat, P. M., and Bunch, M. E., eds. *Interdisciplinary Topics in Gerontology*. Vol. 6, *Age, Work, and Automation*. Basel: S. Karger, 1970. Pp. 69. $6.50.

Peterfalir, Jean-Michel. *Introduction a la psycholinguistique*. Paris: Presses Universitaires, 1970. Pp. 159. Fr 12.

Phillips, Gerald M.; Dunham, Robert E.; Brubaker, Robert; and Butt, David. *The Development of Oral Communication in the Classroom*. Indianapolis: Bobbs-Merrill, 1970. Pp. 211. $6.50.

Phipps, William E. *Was Jesus Married? The Distortion of Sexuality in the Christian Tradition*. New York: Harper & Row, 1970. Pp. xii, 239. $5.95.

Pinner, Edward L., Jr. *A First Group Psychotherapy Book*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xiii, 204. $8.00.

continued on page 406

American Journal of Psychology
1971, Vol. 84, No. 3

TRIAL-AND-ERROR VERSUS ERRORLESS LEARNING: TRAINING, TRANSFER, AND STRESS

Dirk C. Prather
Arizona State University

Ninety-six student pilots were trained on a range-estimation task, either by a trial-and-error method or by an errorless method. After eight training trials, performance under the two methods was statistically equal. In subsequent performance, the trial-and-error subjects were superior on transfer ($p < .05$) and under stress ($p < .01$). Under a combination of transfer and stress, there was no significant difference between the two groups.

Education has traditionally used the trial-and-error method, to teach nearly every type of task. This usually means a student makes a large variety of responses to a single stimulus situation. Through either extrinsic or intrinsic feedback, he slowly narrows down the responses to the most appropriate ones in that particular stimulus setting. The process can lead to frustration and anxiety or to pleasure and self-satisfaction, depending upon the outcomes of the current and past trials. It may also involve the necessary extinction of unwanted responses learned concomitantly with the correct response.

If trial-and-error learning creates problems for the learner in the acquisition of appropriate responses, it would seem that guiding him initially to the correct response in a given stimulus situation should reduce some of the difficulties. If the learner never makes errors in learning a task, many of the uncontrolled, random contingencies of trial-and-error learning may be reduced or eliminated.

The errorless approach is characterized by two types of similar training methods: those that follow the techniques developed by Terrace (1963) and those that train subjects by means of strong prompts. The end result of these two methods is approximately the same in that they both tend to minimize the student's errors. Most of these errorless methods present the subject with some sort of feedback after the response has been made. Quite possibly it is this feedback that the student uses to shape the response, rather than the prompt preceding the response. The present experiment attempted to train the subjects by a method that made error commission during training practically impossible and yet gave the subject no feedback on his performance.

Most of the pertinent experiments cited in the reviews of the literature by Lau (1966) and Prather (1969) indicate that on simple learning tasks, errorless learning should be more efficient when training is restricted to a few trials. The present experiment tested performance on a perceptual skill after four and eight training trials to obtain a measure of efficiency. When transfer is the variable, the literature indicates that trial-and-error learning produces greater transfer of training. Further, most of the experiments on stress have compared performance on a task before and during stress. The present study compared both transfer and performance during stress after training by the trial-and-error and errorless methods.

This experiment was pointed toward flight training in the United States Air Force, as the task involved was one that is similar to flying, and student pilots were used as subjects. The air force is interested in the efficiency of learning because, among other factors, flight training is very expensive. Education in general faces the same problem, of the increasing cost per student hour of instruction. One necessary question, then, relates to the efficiency of any training method. By use of one method will learning take place in fewer trials than by another?

All education and training, no matter where they take place or for which reasons, are pointed toward the transfer of skills and abilities to new stimulus situations. The larger the amount of transfer, the better the training method. So a second question is, When training is completed, will transfer to the real world be better under one method than the other?

In flying, stress is a constant problem in the combat environment and an intermittent one (emergency conditions) in the training situation. There is evidence to show (Harleston, Smith, and Arey, 1965) that the learner's emotional response to stress changes his performance in solving a problem. So a third important question is, Under stress in both the original training situation and a transfer setting, will performance be facilitated more by one training method than the other?

METHOD

—Subjects—Ninety-six male student pilots at Luke Air Force Base were randomly assigned to one of six groups. The age range of the subjects was from 21 to 44, with a mean age of 25.9 years. Every subject was a moderately experienced pilot who had received his wings in a USAF flight-training program.

None of the subjects was experienced in the experimental task. Every subject had passed a stringent physical examination, which included eye and depth-perception tests, within the year preceding the experiment. Some subjects had minor corrections of their eyes to 20/20 vision. All these subjects wore their eyeglasses during the trials. Thirty of the subjects were pilots of the West German Air Force. The remainder of the subjects were USAF pilots. Equal numbers of German pilots were randomly placed in each group. This was the only variable that was controlled.

—Task—Each subject was trained to select a discrete point at which he perceived a target to be a preselected distance from him. The task was not a simple S+ or S− discrimination, but one along a continuum. This skill is much like the ones a pilot must exhibit when flying. The writer, in a preliminary study, found the task suitable as a test of the questions outlined above.

The task required range estimation on simulated strafing runs. Strafing in an actual practice situation usually involves the pilots firing 20-mm cannons at a 20 by 20 ft cloth target. One of the skills a pilot must learn is to estimate his range between himself in the aircraft and the target on the ground. When shooting to get the maximum number of hits, the best slant range to open fire is 2,000 ft. The pilot flies his aircraft toward the target, from approximately 8,000 ft of range, at a closure rate of 800 ft/sec.

A 4-in replica of the actual target was filmed against a plain white backdrop to remove all stimuli used for range estimation except relative target size. The point from which the movie, 16-mm black and white, was started was measured and calculated to give the same relative size and appearance as the actual target on the gunnery range. The camera was moved toward the simulated target at a speed to approximate a 800 ft/sec closure rate over the 8,000 ft of range. Each of the runs was filmed individually. Some were started at less distance, and on several the cameraman operated the camera for 1–2 sec before the run was started. These two actions were taken to vary the number of seconds to correct open-fire point in order to keep the subjects from merely estimating range by means of elapsed time.

The transfer target was a photograph of a MIG-21, a fighter aircraft currently being flown by the North Vietnamese Air Force. The MIG-21 is actually 58 ft long. The photograph was scaled to appear to be the correct size at the various ranges. The procedure for filming the runs on the transfer target was the same as for the training target.

—Apparatus—The stimulus film was projected by a Kodak 16-mm Photo-Optical Data Analyzer onto a standard high-reflection screen.

The subject held a simulated trigger in his right hand. This switch illuminated a light on the experimenter's control box. The experimenter controlled a cue light, which was located directly under the projected image, and a mechanism that delivered a 60-volt shock to the subject's wrists. The experimenter used a .1-sec stopwatch to time each trial.

—Procedure—The sequence of the movie was 14 runs on the training target, 3 runs on the transfer target, and 3 additional runs on the training target. Approximately 5 sec of blank film was inserted between each trial to give the experimenter time to record the stopwatch readings and give feedback when appropriate.

To designate the six groups (TT, ET, TS, ES, TTS, and ETS), the following system of letter abbreviations was used: T, as the initial letter, indicates those subjects trained by the trial-and-error method. If the first letter is an E, the subjects of that group were trained by the errorless method. If the second letter is a T, that group was tested under transfer conditions. The letter S, as either the second or third letter of the designator, indicates that the subjects in the group were tested under stress after training. Group ETS, therefore, was trained under the errorless method, tested on the transfer stimulus, and had a performance-contingent electric shock.

The subjects were seen individually in an academic classroom at Luke AFB. Each subject was asked to sit in a chair 25 ft from the screen and 4 ft to the left rear of the experimenter. When he was seated, the experimeter read the instructions to him.

First, all subjects got this sequence with the *training* target: trained for 4 trials, tested on 3 trials, trained for 4 trials, and tested on 3 trials, for a total of 14 trials.

The trial-and-error subjects were required to press the trigger button when they estimated that the target was at the correct open-fire range. The difference between the correct and actual elapsed time was converted to actual range at open fire. This range was given as feedback to the subject by a verbal statement from the experimenter (e.g., '2400 feet').

The errorless subjects were required to press the trigger when the green cue light came on. This light was illuminated by the experimenter at precisely the correct open-fire range. If the subject began to anticipate the cue light, the experimenter would ask him not to respond until it was lit. This procedure enabled the subject to be trained without response error and yet perform the same motor task as the trial-and-error subjects.

The test trials were identical for both treatments. The experimenter recorded the elapsed time on each trial but did not give any feedback or turn on the cue light. Eight training trials were given, since the pilot study revealed performance should then be statistically equal for both methods.

Next, the subjects were put under conditions of *stress*, *transfer*, or *transfer and stress*. Groups TS and ES saw three more runs on the training stimulus after the shock electrodes were connected. The subjects did not receive any verbal feedback or cue light during these trials, but they did receive an electric shock if they were more than 300 ft away from the desired open-fire range. Each subject received a demonstration shock before these trials.

Groups TT and ET viewed three runs on the transfer stimulus, the MIG-21. These subjects did not receive any verbal feedback or cue light during these trials; however, they were told that they would be given their results at the conclusion of the experiment.

Groups TTS and ETS observed the same three runs on the transfer stimulus. They did not receive any verbal feedback or cue light during these trials, but they did receive an electric shock if they were more than 300 ft off desired range. Each subject received a demonstration shock before the transfer trials.

The experimenter's assistant sat to the left of the subject. The assistant took the subject's pulse for 30 sec at the beginning of the first training trial, at the beginning of the second set of test trials, and immediately after the demonstration shock. If the subject was a member of group TT or ET, where no shock was given, his pulse was taken at the start of the first transfer trial. Training or testing was not stopped while his pulse was being taken; the movie continued to run. The subject did not need his left hand to press the trigger button (a task that is always done with the right hand in fighter aircraft). This pulse-taking procedure did not appear to interfere with training or testing.

RESULTS

The performance data refer to the amount of error in feet from the desired open-fire point, regardless of direction of error. A low score indicates better performance than a high score. The heart-rate data reflect heart rate at the time the distance measures were taken. The data were punched on cards and analyzed by a CDC 3400 computer. All data were compared by *F* test by means of one-way analyses of variance (see Table 1).

Table 1. Analysis of variance

Source	*df*	*MS*	*F*
After four learning trials			
Groups	1	115,426.0	1.59*
Error	94	72,775.5	
After eight learning trials			
Groups	1	142.1	.003
Error	94	52,823.0	
Performance during transfer			
Groups	1	1,626,170	5.47**
Error	30	297,450	
Performance on the training stimulus during stress			
Groups	1	834,244	13.67***
Error	30	61,034	
Performance on the transfer stimulus during stress			
Groups	1	152,159	.55
Error	30	276,510	
Heart rate during training			
Groups	1	662	4.30**
Error	94	154	

*$p < .25$. **$p < .05$. ***$p < .01$.

Training

The results on efficiency after four learning trials were that the mean distance of error was 478 ft for the trial-and-error group and 608 ft for the errorless group. The differences between groups were very marginally significant at the .25 level ($F = 1.59$, *df* 1, 94). The results after eight training trials indicate that the two training groups were almost identical, with a mean error of 418 ft for the trial-and-error group and 416 ft for the errorless group.

Transfer

On the transfer variable, the mean error for group TT was 664 ft; for group ET, 1,115 ft. The difference was significant at the .05 level ($F = 5.47$, *df* 1, 30).

Stress

The results on performance, under stress, on the training stimulus showed a mean error for group TS of 427 ft and of 750 ft for group ES. The difference was significant at the .01 level ($F = 13.67$, *df* 1, 30). Under stress in the transfer situation, the mean error for group TTS was 1,080 ft and 1,218 ft for group ETS. The difference was not significant ($F = .55$, *df* 1, 30).

To demonstrate that stress had been induced, and its effects on performance, groups TT and ET were compared against groups TTS and ETS. The mean error for groups TT and ET was 889 ft, and 1,149 ft for groups TTS and ETS. The difference was significant at the .10 level ($F = 3.51$, *df* 1, 62). These four groups were also compared on heart rate during their performance trials on the transfer stimulus, the third heart-rate measure. The mean heart rate for groups TT and ET was 109; groups TTS and ETS had a mean heart rate of 114. This difference was also significant at the .10 level ($F = 3.07$, *df* 1, 62). These figures show a decrement in performance and a rise in heart rate with the introduction of a performance-contingent shock.

To further demonstrate that stress had been induced and to analyze its effects on performance, groups TS and ES were compared (by a repeated-measures design) on their second training performance and their performance after the shock. This was done by taking measures of performance on the same stimulus just before and during the shock-contingent period. Before the introduction of the electrical shock, groups TS and ES had a mean error of 445 ft. They had a mean error of 588 ft after the performance-contingent shock had been introduced. The difference was significant at the .01 level ($F = 7.80$, $df\,1, 30$). These two groups had a mean heart rate of 108 before shock and 113 after shock. This difference was also significant at the .01 level ($F = 18.40$, $df\,1, 30$).

Interestingly, even when no shock had been given, the errorless group had a significantly higher heart rate after training than did the trial-and-error group. The mean heart rate on the second measure for the errorless group was 110.1 and 104.8 for the trial-and-error group. The difference was significant at the .05 level ($F = 4.30$, $df\,1, 94$). The heart-rate means for both groups, based on measurement during quiescence, before the experiment were 73 beats per minute, according to data in each individual's last physical-examination record sheet.

DISCUSSION

When the literature was surveyed before this experiment, it was noted that most of the experiments involved simple discriminations such as left-right, red-green, circle-ellipse, and so on. Range estimation with few cues on a target closing at 800 ft/sec is certainly more complex, and complex discriminations along a continuum are indeed made by humans many times during the normal day, especially by pilots of high-speed aircraft. Only in experimental studies of auditory discrimination were the tasks comparable in difficulty. And when subjects were required to make these complex auditory discriminations, trial-and-error procedures were superior to heavy prompting (Aiken, 1966). Possibly there is a point in the difficulty of discrimination tasks at which trial-and-error methods overtake and pass errorless procedures in efficiency.

Contrary to most of the literature, the errorless method did not teach the experimental task more efficiently than the trial-and-error method. The mean of the trial-and-error group was 16% better on performance than the mean of the errorless group, although the difference could not be considered significant due to the large amount of variance in performance because of the low number of training trials.

The results on efficiency of training may be contrary to the literature because of the difficulty of the task. In simple discrimination learning, the experimenter can easily analyze the behavior required to form the proper discrimination. This may make it easier for him to set up cues that will control the subject's behavior to the desired end, making success by the errorless method in the teaching of simple discriminations attributable to the narrow limits of the behavior sought. Another reason for the results may be that most adults have a long history of trial-and-error learning; they probably have learned to be efficient at this process and may be able to set up their own intrinsic cues in a complex discrimination, cues more efficient than those an experimenter or educator could extrinsically provide. We might also note that some learning theorists (Holland, 1960; Skinner, 1968; Terrace, 1963) postulate that an incorrect response interferes with the learning of the correct response. This theoretical position is not substantiated by the present results.

In agreement with most of the literature, however, trial-and-error training was superior to errorless training when teaching for transfer. These results extend the literature to a more difficult perceptual task. This is probably the most important finding of the study, since the goal of formal education is the transfer of skills and abilities to new stimulus situations.

The finding that subjects trained under the trial-and-error method performed significantly better under stress is very important for the training of pilots. The differences were significant when the training stimulus was used; they were not significant for the transfer stimulus, although the performance of the trial-and-error subjects on the transfer stimulus was 13% better than that of the errorless subjects. The lack of significant findings under the transfer-and-stress condition (groups ETS and TTS) was probably because both groups received the performance-contingent electric shock. With subjects using this information as feedback to improve their responses, performance differences would tend to narrow. There is little data on this question, for most previous experiments have compared performance before and during stress. The advantage of the present method of comparing performance under stress after different kinds of training is that it does allow the findings on stress to be interpreted in a fashion similar to that of most experiments on transfer.

There is ample proof that stress was induced in the present investigation. The significant rise in heart rate and the decrement in performance between measures taken just before the electric shock and during the shock-contingent period both indicate that stress was induced. The differences in heart rate and performance on the transfer stimulus with shock and without are another indication. So is the fact that groups TTS and ETS had a mean heart rate of 114 during the shock-contingent period. Compared with the heart rate of 118 taken on pilots in combat under heavy antiaircraft fire during bombing runs over Hanoi (Plattner, 1967a, 1967b), this is convincing evidence of stress.

The experiment suggests that when prompts are too powerful, they may decrease some of the involvement in the task or make the involvement more passive. The trial-and-error subjects were much more actively involved in the learning situation in that they seemed to treat the task as a game and to display many more verbalizations during the experimental period. The errorless subjects seemed much more passive in the situation.

The significantly higher heart rate displayed by the errorless subjects after training is also an interesting finding on which to speculate. Since both errorless and trial-and-error subjects had the same mean heart rate before training, some variable in the errorless training method seems accountable for the difference after training. The fact that the errorless subjects were aware that test trials were imminent and that they had no knowledge of their previous performance may be important here.

The evidence from this experiment clearly implies that a trial-and-error method of training is superior to an errorless procedure when the task is one of difficult perceptual learning. It may well be that the trial-and-error method should be extensively employed when teaching a perceptual-motor skill such as flying and that the indiscriminate use of errorless procedures for teaching complex tasks is open to question.

Notes

The opinions expressed and the conclusions drawn in this paper are those of the author and do not necessarily reflect the views of the United States Air Force or the Department of Defense. The author's present address is at the United States Air Force Academy. Received for publication July 23, 1970.

References

Aiken, E. G. 1966. Multiple versus single stimulus discrimination learning of active sonar pip shapes under two training formats. *USN PRA Technical Bulletin* STB 67–11.

Harleston, B. W., Smith, M. G., and Arey, D. 1965. Test-anxiety, heart rate and anagram problem solving. *Journal of Personality and Social Psychology* 1:551–557.

Holland, J. G. 1960. Teaching machines: An application of principles from the laboratory. In *Teaching machines and programmed learning*, ed. A. A. Lumsdaine and R. Glaser, pp. 215–228. Washington D.C.: National Education Association.

Lau, A. W. 1966. A comparison of prompting versus feedback in verbal and perceptual learning. *USN PRA Technical Bulletin* STB 67–8.

Plattner, C. M. 1967a. Carrier pilots show little stress in North Vietnam raids. *Aviation Week and Space Technology* 86(17):80–85.

Plattner, C. M. 1967b. Heart strain greater in landing on carrier than bombing. *Aviation Week and Space Technology* 86(11):60–65.

Prather, D. C. 1969. The effects of trial-and-error or errorless training on the efficiency of learning a perceptual-motor skill and performance under transfer and stress. *Dissertation Abstracts International* 20:790.

Skinner, B. F. 1968. *The technology of teaching*. New York. Appleton-Century-Crofts.

Terrace, H. S. 1963. Discrimination learning with and without errors. *Journal of the Experimental Analysis of Behavior* 6:1–27.

American Journal of Psychology
1971, Vol. 84, No. 3

THE SWEETNESS AND PLEASANTNESS OF SUGARS

Howard R. Moskowitz
Pioneering Research Laboratory, U.S. Army Natick Laboratories

Nine experiments were run to assess the relation between sweetness, pleasantness, and the concentrations of 43 sugars. A power function adequately related the sweetness, S, of most sugars to their concentration, C: $S = kC^n$, where the intercept k is a measure of relative sweetness. The average exponent is about 1.3, and the results suggest that the large range of relative sweetness across sugars is correlated with differences in molecular structure. The pleasantness of sugars was not monotonic with concentration but systematically departed from linearity at high concentrations. Pleasantness as a function of sweetness was roughly linear in log-log coordinates, with a slope between .3 and .5.

A number of studies (Meiselman, 1968; Moskowitz, 1970a, 1970b; Stevens, 1969b) have suggested that the sweetness, S, of sucrose grows in magnitude as a power function of concentration, C: $S = kC^n$. In log-log coordinates this relation between the perceptual continuum of sweetness and the objectively measured continuum of concentration is rectified into a straight line, with a slope of n and an intercept of k. The exponent is particularly important for characterizing the dynamic growth of sweetness because it indicates how sucrose concentrations are transformed into reported sweetness. If n exceeds 1.0, then sweetness accelerates with concentration, whereas if n is less than 1.0 then sweetness decelerates with concentration. An exponent of 1.0 indicates that the growth of sweetness is linear with the growth of concentration.

Estimates of the sweetness exponent of sucrose obtained from various laboratories and procedures give conflicting results as to whether sweetness is a negatively or positively accelerating function of the sweetener, Kocher and Fisher (1969) and Feallock (1965) reported exponents of approximately .5; both studies used a flow system to present the taste stimulus. Meiselman (1971) has demonstrated that the procedure of presentation can drastically affect the sweetness exponent, with the flow method typically giving lower values than the sip procedure. Experimenters who used the sip procedure have reported higher exponents.

Those obtaining values below 1.0 include Gregson and Russell (1965), who got an *n* about .62, and Meiselman (1968), who got an *n* about .70. Stone and Oliver (1969) reported an exponent around 1.0, Stevens (1969b) an exponent around 1.3, Moskowitz (1968, 1970a, 1970b) exponents between 1.3 and 1.6, and Ekman and Åkesson (1964) an exponent around 1.5.

Fewer studies have been published on the measurement of the 'pleasantness,' or affective dimension, of sugar. Engel (1928) showed that at intermediate concentrations a large proportion of his observers reported that sucrose was pleasant, but at higher concentrations many of them noted that sucrose became unacceptable. In a choice experiment, Chappell (1953) instructed observers to select the most acceptable concentration of sucrose from an array of eight solutions and found that, on the average, observers selected 25% sucrose as the most acceptable. In a further experiment, she showed that the type of sugar also influences acceptability, with the descending order of preference being sucrose, lactose, glucose, and maltose. Unfortunately, she did not obtain numerical estimates of the degree of sugar pleasantness. Ekman and Åkesson (1964) and Kocher and Fisher (1969) reported some pilot studies that did assess the magnitude of pleasantness. Ekman and Åkesson showed that there are large individual differences in the acceptability of sucrose solutions, although the group results suggest that sugar pleasantness is an inverted-V function of sugar sweetness. Kocher and Fisher also measured pleasantness as a function of sweetness, but they found that the function was monotonic rather than shaped like an inverted V.

The present study assesses the relation between sweetness, pleasantness, and concentration. 'Magnitude estimation,' or free number matching, was used to obtain the functional relation between the concentration and the perceptual and affective dimensions. The method of direct scaling of taste magnitude has also been used successfully by several of the above-mentioned experimenters to determine the relation between sweetness and concentration.

METHOD

—Stimuli and procedure—In seven experiments a total of 32 sugars were scaled for both sweetness and pleasantness. The judgments of each dimension were obtained on alternate days from different groups of observers selected from a pool of 63 volunteers. In the eighth experiment 12 observers judged the sweetness of 10 additional sugars, and in the last experiment 13 observers judged the sweetness of glucose and tagatose. Table 1 lists the concentrations of the sugars, as well as the conditions of the experiments.

All sugars were reagent grade (Sigma Chemical or Nutritional Biochemicals), and solutions were prepared by mixing measured amounts of sugar with distilled water until the appropriate molarity was obtained. In most of the experiments the concentration of the stimuli formed a geometric sequence, with the ratio of the progression being 2:1. The sugar solutions were presented in small paper (¾-oz soufflé) cups, each holding 5–7 ml of stimulus, which were arranged in irregular order on a lazy susan. The observer sampled the stimuli and recorded his estimates on the answer sheet provided. The solutions were served at room temperature (approximately 21 deg C ± 1 deg), but when not in use they were refrigerated to prevent decomposition. Before each experiment the solutions were left under refrigeration for 48 hours in order to allow for a mutarotational equilibrium to be established.

The observers were given these written instructions:

> In front of you is a series of solutions in paper cups. Your task is to tell how sweet [*or* pleasant] they seem by assigning numbers to them. If the second stimulus is eleven times as sweet [pleasant], then assign it a number eleven times as large; if it seems one-nineteenth as sweet [pleasant], then assign it a number one-nineteenth as large, and so forth. Use fractions, decimals, and whole numbers, but make each assignment proportional to sweetness [pleasantness]. If the solution has no sweetness [*or* If it seems either neutral or unpleasant], then assign it a '0.' All other assignments, whether sweetness or pleasantness judgments, are to be positive numbers.

Observers who had never participated before in experiments in direct scaling of sensory magnitude were trained in the use of number matches by being shown different line lengths and circular areas and given instructions to match numbers to these stimuli in proportion to apparent length and apparent area. Observers continued to assign numbers to the training stimuli until they understood the technique of number matching.

There were no time limitations on the experiment, except that observers were told to rinse their mouths fully between trials in order to remove traces of a preceding stimuli. Most observers proceeded at rapid enough a pace that they completed the sessions within 15 to 20 min. Experiment IX required about 5 to 10 min.

—Analysis of the magnitude estimates—Since the magnitude estimates of the observers were not restricted to a given modulus (scale unit), an initial normalization was needed in order to standardize the modulus of each observer in all of the experiments. For each observer the normalization consisted of multiplying each of his judgments of the five standard glucose stimuli by a factor that made the geometric mean of the five estimates equal 10.0. The remaining judgments of the same observer to the other sugars in the experiment were also multiplied by the same normalizing factor. Since these same five glucose concentrations were presented to each observer in every experiment, they served as an implicit standard across experiments and observers. By this normalization procedure the magnitudes of numbers used by different observers in the nine experiments were roughly equated. Because the estimates of each observer in a single experiment were each multiplied by the

Table 1. Stimulus concentrations and experimental conditions

Sugar	(Molecular weight)	Molar concentration 1	2	3	4	5	6	7
Experiment I (32 observers judging sweetness, 36 pleasantness)								
Glucose	(180)	.125	.25	.5	1.0	2.0		
Dulcitol	(182)	.125	.25					
Glycerol	(92)	.125	.25	.5	1.0	2.0	4.0	8.0
Ribose	(150)	.125	.25	.5	1.0	2.0		
D-Xylose	(150)	.125	.25	.5	1.0	2.0		
Experiment II (36 observers judging sweetness, 36 pleasantness)								
Glucose	(180)	.125	.25	.5	1.0	2.0		
L-Arabinose	(150)	.125	.25	.5	1.0	2.0		
Fructose	(180)	.125	.25	.5	1.0	2.0	3.4	
Sorbose	(180)	.125	.25	.5	1.0	2.0		
Experiment III (31 observers judging sweetness, 31 pleasantness)								
Glucose	(180)	.125	.25	.5	1.0	2.0		
Galactose	(180)	.125	.25	.5	1.0	2.0	3.0	
Inositol	(180)	.125	.25	.5				
Mannitol	(182)	.125	.25	.5	1.0			
Mannose	(180)	.125	.25	.5	1.0	2.0	3.0	
Experiment IV (34 observers judging sweetness, 34 pleasantness)								
Glucose	(180)	.125	.25	.5	1.0	2.0		
Maltose	(342)	.0625	.125	.25	.5	1.0	2.0	
Sorbitol	(182)	.125	.25	.5	1.0	2.0	3.0	
Sucrose	(342)	.0625	.125	.25	.5	1.0	2.0	2.5
Experiment V (35 observers judging sweetness, 32 pleasantness)[a]								
Glucose	(180)	.125	.25	.5	1.0	2.0		
Lactose	(342)	.0625	.125	.25	.5			
Raffinose $.5H_2O$	(594)	.0625	.125	.25				
Rhamnose	(164)	.125	.25	.5	1.0			
Experiment VI (12 observers judging sweetness, 13 pleasantness)[b]								

Cellobiose	(342)	.0625	.125	.25		
L-Erythritol	(122)	.125	.25	.5	1.0	2.0
Melezitose $.2H_2O$	(540)	.0625	.125	.25	.5	.75
Melibiose $.2H_2O$	(378)	.0625	.125	.25	.5	1.0
Trehalose $.2H_2O$	(378)	.0625	.125	.25	.5	
Experiment VII (12 observers judging sweetness, 13 pleasantness)[c]						
Glucose	(180)	.125	.25	.5	1.0	2.0
α-L-Fucose†	(164)	.125	.25	.5	1.0	2.0
D-Lyxose	(150)	.125	.25	.5	1.0	2.0
Ribitol†	(152)	.125	.25	.5	1.0	2.0
D-Turanose	(342)	.0625	.125	.25	.5	
D-Xylitol	(152)	.125	.25	.5	1.0	2.0
L-Xylose	(150)	.125	.25	.5	1.0	2.0
Experiment VIII (12 observers judging sweetness, each sugar judged in two separate sessions)						
Glucose	(180)	.125	.25	.5	1.0	2.0
2-Deoxy-D-galactose	(164)	.0625	.125	.25		
6-Deoxy-L-galactose†	(164)	.125	.25	.5	1.0	
2-Deoxy-glucose	(164)	.125	.25	.5	1.0	2.0
6-Deoxy-mannose	(164)	.125	.25	.5	1.0	2.0
2-Deoxy-ribose	(134)	.125	.25	.5	1.0	
β-D-Fucose	(164)	.125	.25	.5	1.0	
Glucoheptose	(210)	.125	.25	.5		
Mannoheptulose	(210)	.125	.25	.5		
Perseitol	(212)	.0625	.125	.25	.5	
Sedoheptulose anhydride	(192)	.0625	.125	.25	.5	
Experiment IX (13 observers judging sweetness)						
Glucose	(180)	.125	.25	.5	1.0	2.0
Tagatose	(180)	.25	.5	1.0		

---

† Ribitol and adonitol, and α-L-fucose and 6-deoxy-L-galactose are the same sugars structurally but are marketed under two different labels and were tested as though they were different sugars.

[a] Lactose and raffinose each presented twice, in irregular order.

[b] Glucose presented twice in the series.

[c] Glucose presented twice in the series.

same factor, the ratios of estimates remained unchanged, and therefore the slope (exponent) of the sweetness function was unaffected. However, the variability due to different number sizes was substantially reduced.

The data were then analyzed with a computer program called PSYCHOFIT (Panek and Stevens, 1965), revised by the author. The revised program provided summary statistics on the distribution of the magnitude estimates, as well as least-squares estimates of the exponent and intercept of the power function $S = kC^n$ fitted to the empirical geometric means and medians. PSYCHOFIT also provided a least-squares estimate of the intercept k of the power function when the exponent was fixed at a predetermined value.

The median was used as the measure of central tendency because several observers gave magnitude estimates of '0.' The geometric mean is usually the appropriate mean because it preserves ratios, but here it is an inappropriate statistic because when one or more judgments is '0' the geometric mean itself automatically becomes '0.'

RESULTS AND DISCUSSION

Figures 1–4 present the relation between molar concentration and the median judgments of both sweetness (Figures 1–4) and pleasantness (Figures 1–3). Concentrations that fell below the taste threshold are not

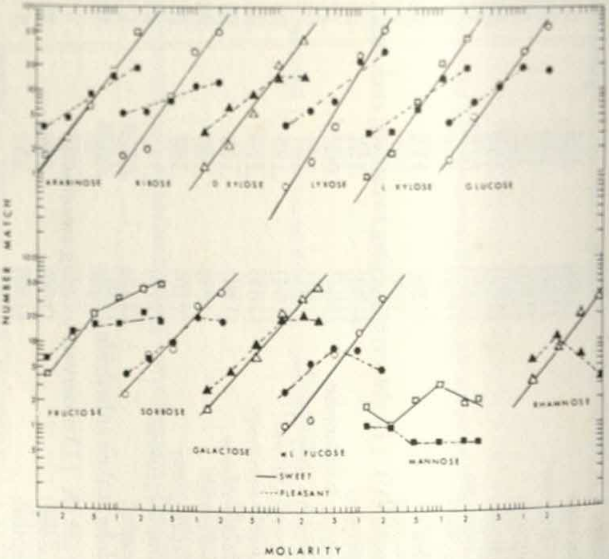


Fig. 1. The sweetness and pleasantness of simple sugars, plotted in log-log coordinates. Straight lines suggest power functions relating the subjective continua to the objective continuum of concentration.

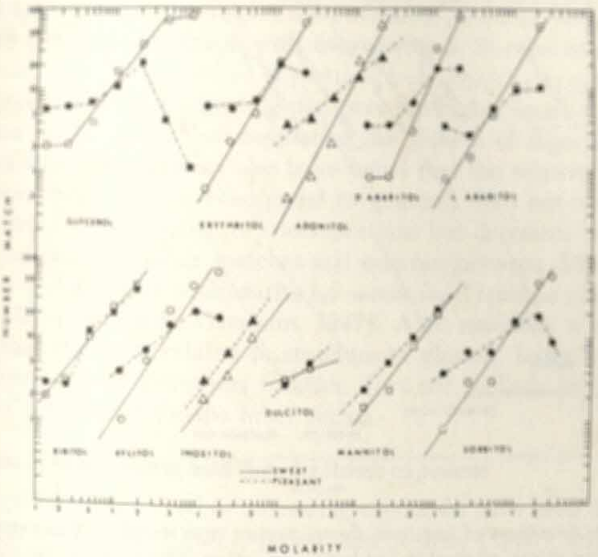


Fig. 2. The sweetness and pleasantness of sugar alcohols, plotted in log-log coordinates

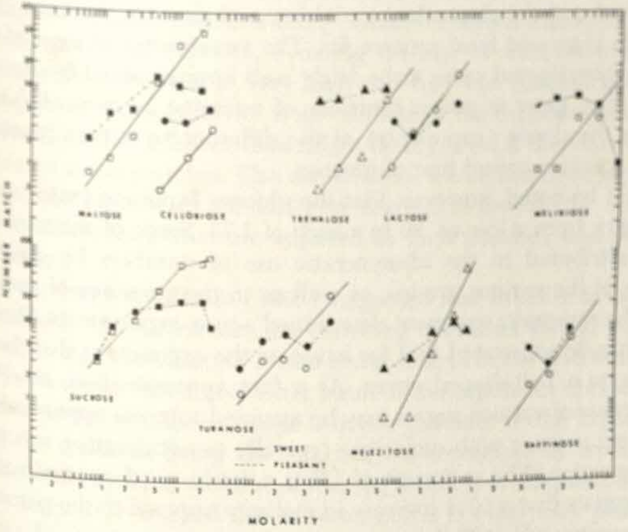


Fig. 3. The sweetness and pleasantness of oligosaccharides (heavier sugars) plotted in log-log coordinates

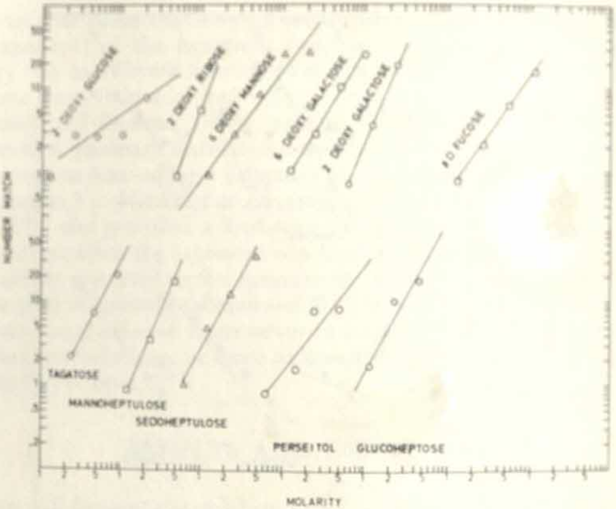


Fig. 4. The sweetness of tagatose, deoxy sugars, and simple sugars containing seven carbon atoms, plotted in log-log coordinates. The ordinate for 2-deoxygalactose has been displaced upward by one log unit.

presented in the figures. The coordinates are logarithmic for both the independent and the dependent variables, and the straight lines fitted to the data represent least-squares fits. The sweetnesses of most of the 43 sugars investigated seem to be fairly well approximated by straight lines and thus grow as power functions of molarity. In most of the experiments the slopes (exponents) of the different sweetness functions (Table 2) cluster around that of glucose.

It should be noted, however, that the glucose exponent varies across experiments, from a low of .90 to a high of 1.73. Some of the variation may be attributed to the idiosyncratic use of numbers by different observers in the various groups, as well as to the presence of random error in the sweetness exponent determined across experiments. Stevens (1969a) has demonstrated that for loudness the exponent is distributed according to a bell-shaped curve. As a first approximation, however, the sweetness of various sugars may be assumed to grow approximately in the same manner with molarity—typically as accelerating functions of concentration. The exponent of 1.3 is a fairly good approximation, and it suggests that a 10:1 increase in molarity appears to the perceiver to be approximately a 20:1 increase in sweetness.

Several small but systematic departures from linearity occur for different sugars. For some sugars the sweetness functions become ogival instead of remaining linear with concentration. Sucrose and fructose, however, show more marked departures from linearity at medium and high concentrations, suggesting that over the higher levels they do not conform to the power law regulating the growth of sugar sweetness. Several other investigators also have noted that the relative sweetness of sucrose and fructose (compared to glucose) does not remain constant across the full range of concentrations but decreases at the high end. Both direct number matches and matches between different solutions have shown this nonlinearity for sucrose and fructose (Moskowitz, 1970a; Stevens, 1969b; Cameron, 1947). Also, mannose, a six-carbon sugar that is closely related in structure to glucose, tastes both bitter and sweet, but its sweetness function does not conform to the power relation determined for the other sugars.

Regression in matching and the sweetness exponent

The problems involved in measuring exponents have been dealt with at length by Stevens (1957), Indow and Stevens (1966), and Stevens and Greenbaum (1966). Direct matching procedures such as magnitude estimation are subject to several biases, one of which is the so-called regression effect, which shows up clearly when observers are instructed to match stimuli from one continuum (here, numbers) to another continuum (here, sugar concentration). Observers often constrict the range of their numerical matches, avoiding the use of extreme numbers to represent the sweetness of very high and very low concentrations. One consequence of this behavior is an estimate of the exponent relating the two continua that is lower than might be expected if the observer were not biased in his matches. The constriction, which has been labeled 'regression' because of its similarity to 'regression toward the mean,' applies to all continua that are adjusted in such matches, and not solely to numbers.

Stevens and Greenbaum (1966) suggested that in order to cancel out the bias due to constriction, the matching operation should be turned around, so that the observer is also given control of the criterion modality. They assume that a geometric mean of the exponents obtained from matches in the two directions is a better estimate of the exponent relating the continua being studied. But the technique of continuously altering concentration without producing artifacts of adaptation has not been successfully worked out in taste, nor is it clear why performing the matches in the two directions would eliminate the regression effect, as they suggest it would.

Table 2. Parameters of the power function $S = kC^n$ relating sweetness and concentration

Sugar	Pearson r	n Least-squares exponent	Standard error of exponent	k Molarity intercept	k^1 Percentage intercept	Standard error of log intercept[a]
Experiment I						
Glucose	.996	1.31	.106	29.21	.65	.000
Dulcitol	1.000	.48	.287	4.43	1.51	.070
Glycerol	.972	1.14	.082	11.53	.93	.036
Ribose	.975	1.35	.107	20.61	.52	.059
D-Xylose	.992	1.31	.097	16.71	.48	.048
Experiment II						
Glucose	.992	1.09	.089	22.18	.94	.000
Arabinose	.993	1.22	.108	17.71	.84	.030
Fructose	.964	.99	.069	39.92	2.26	.028
Sorbose	.981	1.04	.062	23.46	1.15	.031
Experiment III						
Glucose	.980	1.25	.137	24.83	.66	.000
Galactose	.993	1.16	.135	17.65	.61	.049
Inositol	.999	1.35	.203	22.34	.45	.050
Mannitol	.990	1.26	.126	20.19	.52	.042
Experiment IV						
Glucose	.989	1.31	.098	27.33	.61	.000
Maltose	.987	1.30	.124	24.79	.25	.046
Sorbitol	.984	1.37	.110	14.81	.28	.043
Sucrose	.983	1.03	.087	59.38	1.56	.030
Experiment V						
Glucose	.984	1.05	.215	23.82	1.13	.000
Lactose	.999	1.21	.222	35.28	.48	.081
Raffinose	.999	1.02	.333	24.74	.38	.102
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
D-Arabitol	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

		2.30	.430	21.06	.04	.091
L-Arabitol	.844	1.37	.335	13.75	.33	.084
Cellobiose	.845	1.31	.324	18.54	.18	.182
Erythritol	.994	1.69	.460	20.81	.30	.078
Melibiose	.948	1.17	.214	13.46	.19	.137
Melezitose	.968	2.35	.492	99.14	.01	.090
Trehalose	.964	1.27	.345	18.57	.18	.174
Experiment VII						
Glucose	.988	1.70	.275	27.01	.26	.000
α-L-Fucose	.974	1.37	.338	14.87	.32	.101
Lyxose	.987	1.63	.202	17.78	.21	.079
Ribitol	.986	1.50	.267	21.01	.35	.029
Turanose	.826	.84	.360	15.02	.77	.094
Xylitol	.960	1.48	.271	33.54	.60	.044
L-Xylose	.990	1.46	.191	17.97	.36	.093
Experiment VIII						
Glucose	.937	1.39	.197	26.20	.47	.000
2-Deoxy-galactose	.999	1.82	.377	6.15	.62	.104
6-Deoxy-galactose	.998	1.55	.273	26.27	.34	.083
6-Deoxy-glucose	.977	1.28	.204	16.50	.46	.090
2-Deoxy-mannose	.283	.65	.209	4.94	.80	.134
2-Deoxy-ribose	1.000	2.61	.286	5.38	.006	.108
β-D-Fucose	.999	1.45	.260	16.57	.29	.109
Glucoheptose	.953	1.70	.239	67.97	.42	.100
Mannoheptulose	.999	2.17	.295	78.10	.12	.092
Perseitol	.948	1.26	.199	26.03	.55	.085
Sedoheptulose	.973	1.67	.173	125.02	.90	.082
Experiment IX						
Glucose	.988	.90	.231	18.66	1.38	.000
Tagatose	.999	1.63	.332	21.73	.20	.132

Note: Standard error of the intercepts was computed after each individual's slope was forced to equal the grand slope. Since the glucose responses were normalized to the same value for each subject, the standard error of the intercept equals 0.0 for glucose. The names for the sugars in this table are somewhat abbreviated to correspond to the figures; for full names, see Table 1.

However, there is another matching paradigm that may effectively reduce the regression and thus permit a more accurate determination of the sweetness exponent. If the observer is instructed to match a third continuum (say, sound pressure, P, which is experienced as loudness) to both the concentration of sugar, C, and to the magnitude of numbers, N; then two separate relations may be obtained, both with sound pressure, P, as the common modality;

$$P = C^{a} \text{ (pressure-concentration or loudness-sweetness match),} \quad [1]$$

$$P = N^{b} \text{ (pressure-number or loudness-number match), or} \quad [2]$$

$$N = C^{a/b} \text{ (number-concentration relation obtained by eliminating } P\text{)}. \quad [3]$$

The relation between number and concentration (Equation 3) is a derived function, obtained from two empirically determined matching functions in which both the continua of number and of concentration serve as the stimulus modalities while pressure is varied to match each. If the physical ranges of both concentration and number are commensurate, then the constriction that occurs in the range of pressure matches should be identical in both sets of matches.

The use of a third modality (sound pressure) to relate the continua of concentration and number has been reported by Moskowitz (1968), who instructed his observers to match the loudness of a wide band of white noise (500–5,000 Hz) to the sweetness of sucrose. The exponent relating pressure, P, and sucrose concentration, C, turned out to be approximately 2.0, suggesting that loudness grows as the square of sugar concentration. A good estimate of the pressure-number exponent, when pressure is adjusted to match number, is approximately .75.[1] From Equations 1–3 the exponent relating concentration to number is thus approximately $.75 \times 2.0 = 1.5$, which is higher than the 1.3 exponent obtained in the present studies. The higher exponent suggests that sugar sweetness is an even more positively accelerated function of concentration than it appears to be from number matches.

The relative sweetness of sugars

In spite of structural similarities and comparable rates of sweetness growth with concentration, sugars differ in their sweetness. Relative sweetness of sugars can be defined by the ratio of numerical estimates of sweetness of two sugars at a common concentration, although if the sugars grow in sweetness at different rates, then their relative sweetness will continuously change across concentrations. Table 2 presents the sweetness level of each sugar at a 1-molar concentration, which means that the assessment of sweetness is made upon the molecule of sugar, and at a 1% concentration, which means that the assessment is made on the *weight* of sugar in solution without regard to the size of the sugar molecule. Glucose can be assigned a reference value of 1.0 in both cases, so that the relative sweetness can be expressed as fractions or multiples of glucose sweetness. But because the sweetness functions were not adjusted to be parallel in log-log coordinates, the estimates of sweetness in Table 2 pertain only to 1-molar and 1% concentrations respectively and cannot be generalized to other concentrations.

For convenience in assessing the sweetnesses of different sugars, the sweetness exponent was then fixed at a representative value of 1.3. When the exponents (slopes) of the sugar functions are thus equalized, they become parallel lines in log-log coordinates and their logarithmic distance (i.e., their numerical ratio) does not change over the range of concentration. The method of fixing the exponent permits a single number to characterize the relative sweetnesses of different sugars. Figure 5 shows the relative sweetnesses of various sugars in parentheses, in which the left-hand number represents sweetness at 1 molar and the right-hand number represents sweetness at 1%.

Figure 5 represents a 'tree structure' of sugars, with sugars at the same level (e.g., C_5 or C_6) being closely related and derived from the sugars directly below them. For example, the six-carbon sugar galactose (C_6) is derived from the five-carbon pentose sugar, lyxose, which in turn is derived from a lower sugar containing four carbon atoms. The branches to one side and tilted 45 deg are both sugar alcohols (e.g., dulcitol) produced by changing the aldehyde group CHO into the reduced form CH_2OH, and deoxy sugars obtained by removing an oxygen atom from the parent sugar (e.g., fucose and the deoxy galactoses). Glucose has been assigned a reference value of 1.0, and all sugars have been measured relative to the sweetness of glucose.

The compound sugars occupy the upper portion of Figure 5. These di- and trisaccharides are obtained when two or three simple sugars (monosaccharides) are joined together in one or another manner. The combination may be homogeneous between two glucose molecules (e.g., maltose, cellobiose, and turanose), or it may be heterogeneous between dissimilar molecules (e.g., glucose and either fructose or galactose). The union may be between different carbons of the sugar ring (e.g., at the carbon atoms numbered 4 or 6), and in addition the bond can be either of the α or β forms, which differ only stereochemically. The parentheses indicate what link binds together the constituent monosaccharides.

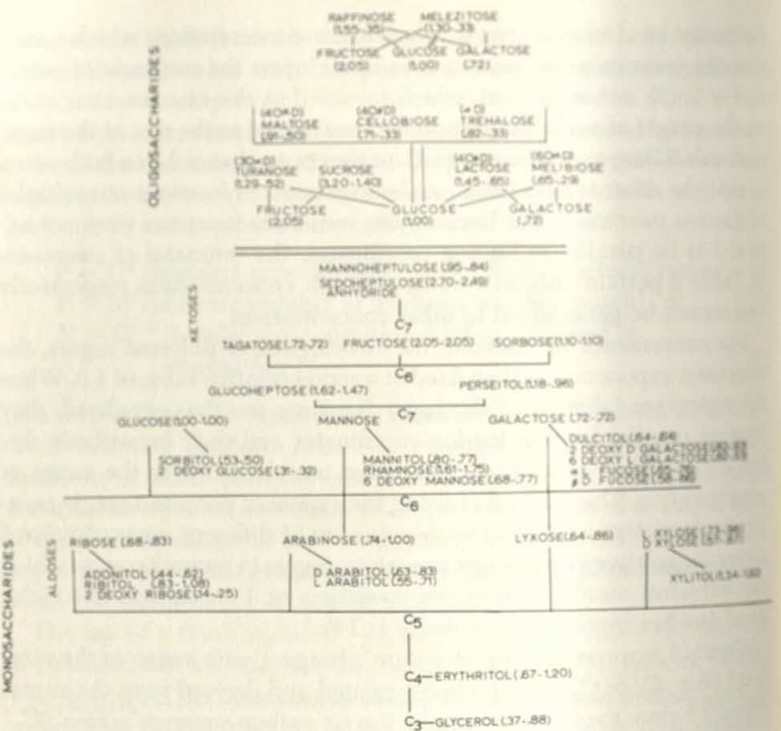


Fig. 5. Relative sweetness as a function of chemical structure. The numbers in parentheses represent relative sweetness at 1-molar (left-hand number) and 1% (right-hand number) concentrations. Glucose has been assigned a reference value of 1.0. Vertical lines indicate that the higher sugar has been derived from the lower one; lines tilted 45 deg indicate that the sugar is either a sugar alcohol or a deoxy derivate of the parent sugar.

Several systematic changes in relative sweetness occur when the structure of the sugar is changed or the combination of sugars is altered:

1. Per molecule the five-carbon sugars are less sweet than the six-carbon ones, although per gram the ordering of sweetness is changed.
2. The primary effect of reducing an aldose sugar to its sugar-alcohol derivative is to reduce sweetness.
3. The ketose sugars, having the ketone group (C=O), are usually sweeter than the aldoses having the aldehyde group (CHO).
4. The presence of the furanose form of fructose increases sweetness considerably (for fructose and sucrose), but it also appears to lead to a departure from the simple power relation governing sweetness

5. The seven-carbon sugars appear to be about as sweet or somewhat sweeter than the six-carbon sugars.
6. The link between the constituents of an oligosaccharide partially determines its relative sweetness, especially in combinations of two different sugar molecules that form heterogeneous disaccharides. On the other hand, for homogeneous combinations of two glucose molecules the type of linkage does not appear to be as crucial.
7. Sucrose and lactose each appear to be sweeter than their constituents, although sucrose sweetness appears to be more nearly the sum of the individual glucose and fructose components. This increase in sweetness, which suggests partial or full additivity of the component sweetnesses, does not appear to hold for raffinose, turanose, melezitose and melibiose, which are far less sweet than might be expected from the sweetnesses of their constituents.
8. As a first approximation the three homogeneous disaccharides, each of which combines two glucose molecules, appear to be about as sweet as a single molecule of glucose, suggesting that only half of the compound molecule effectively stimulates the taste system.

Other estimates of the relative sweetness of sugars have been reported by Moskowitz (1970a), who has tabulated several values reported in the literature. In that article, sucrose, rather than glucose, was the reference sugar.

Figure 6 presents the relative sweetness of several sugars as a function of the representative exponent chosen for sugar sweetness. The relative sweetnesses of sugars do not change across different representative exponents if the sugars are presented at the same five concentrations as the standard five glucose levels (that is, at molar concentrations of .125, .25, .5, 1.0, and 2.0 respectively). Those sugars presented in Figure 6 were presented in different concentrations, however, and therefore their sweetnesses vary. The functions are plotted in semilogarithmic coordinates and allow for extrapolation beyond the limits of 1.1 and 1.5. The estimates of sweetness thus permit a revaluation of representative sweetness values if the taste exponent for sugars is changed from 1.3.

The pleasantness functions

Judgments of the pleasantness of sugars are indicated in Figures 1–3 by dashed lines connecting the filled markers. Unlike sweetness, the pleasantness functions of sugars do not appear to be monotonic with concentration; rather, they systematically depart from linearity at the extremes of concentration. Several of the sugars first increase in accept-

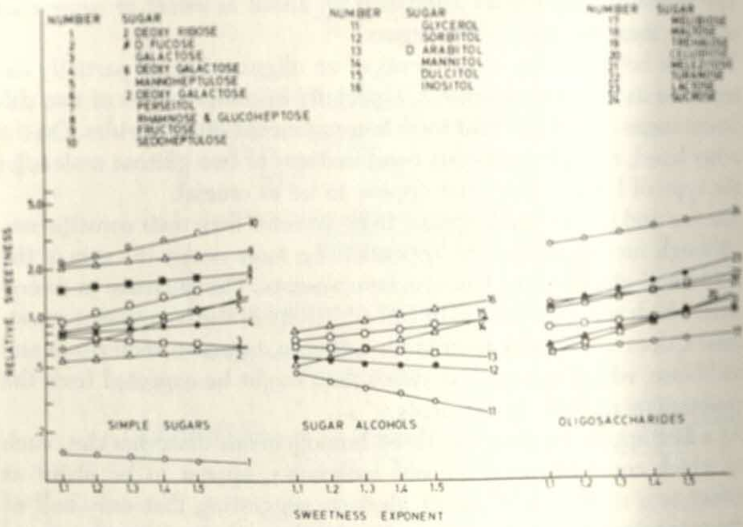


Fig. 6. The relation between representative exponent for sugar sweetness and the relative sweetness of sugars. The sugars presented in this figure were assessed at concentrations different from the standard five levels of glucose, and therefore their relative sweetnesses across different representative exponents do not remain the same compared to glucose.

ability with concentration and then decrease. This break is obvious for most of the sugars except the pentoses, which continue to increase in pleasantness with concentration. Mannose, which is both bitter and sweet, is the least pleasant of the sugars, but the removal of an oxygen atom to obtain rhamnose produces a sugar that is both sweet and pleasant.

Another way of analyzing the relation between sweetness and pleasantness is to consider, for each concentration, the pleasantness judgment as a function of the corresponding sweetness judgment. This plot is shown in Figure 7, which represents the sweetness-pleasantness relation for three families of sugars. The coordinates are log-log, and the straight lines were fitted by a least-squares procedure. The slope relating log pleasantness to log sweetness ranges between .3 and .5, suggesting that the apparent pleasantness of a sugar grows between the cube root and the square root of its apparent sweetness. At high sweetness levels the pleasantness of sugars begins to decrease with subsequent increases in sweetness—a decrease similar to that found by Engel (1928).

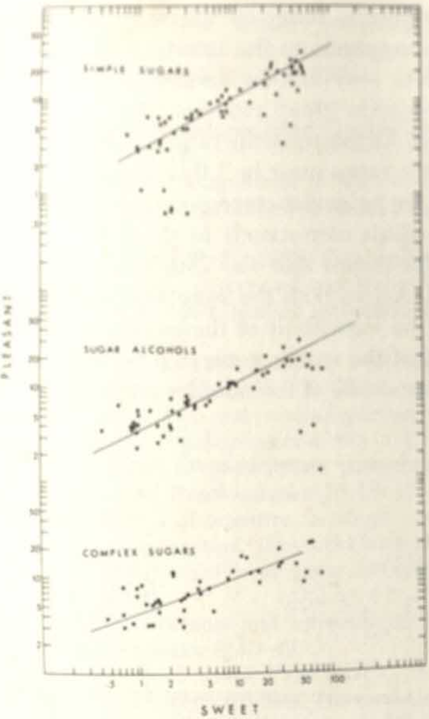


Fig. 7. The relation between pleasantness of sugars and sweetness of sugars. Each point is the median judgment of both sweetness and pleasantness of a single concentration of sugar. The coordinates are log-log, and the slopes range between .3 and .5.

Figure 7 may be used to calculate 'trade off' relations between sweetness, pleasantness, and concentration. For example, one can assess the effect on pleasantness of increasing or decreasing sugar content, as well as measure the change in acceptability occurring when one sugar is substituted for another. In addition, Figures 1–4 and 7 show the range of interconvertibility of sugars that will maintain either sweetness or pleasantness. The results were obtained for aqueous solutions of sugars, so specific food items would have to be investigated to generate similar data for foodstuffs.

Variability in direct number matches

Despite their slope differences, both the sweetness and pleasantness judgments exhibit roughly the same magnitude of variability. Figure

8 shows the interquartile range of the logarithms of the numerical matches, which corresponds to the interquartile ratio of the matches themselves. The data were first normalized to adjust the moduli of individuals to the common group modulus, and the interquartile range was then computed. All interquartile ranges must be 0.0 or greater, and thus all interquartile ratios must be 1.0 or more. An interquartile ratio of 1.0 means that the judgment corresponding to the 75th percentile is the same as that which corresponds to the 25th percentile. An interquartile ratio of 5.0 means that the 75th percentile is five times the 25th, and so forth. Across both the sweetness and pleasantness scales, the variability of the logarithms of the matches tends to remain constant, suggesting that the scatter generated by number matches is proportional to the magnitude of the number match.

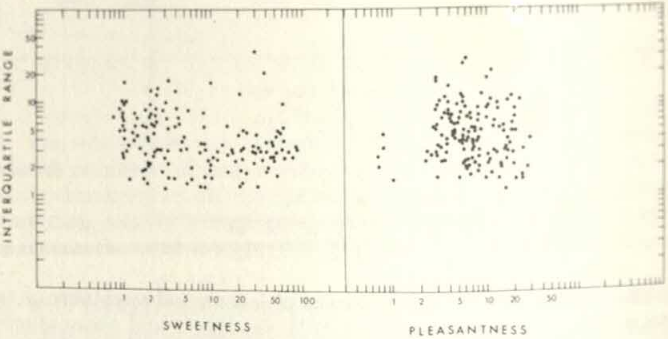


Fig. 8. The variability of magnitude estimates as a function of the magnitude of sweetness and pleasantness. The coordinates are log-log.

Notes

Eugene Zanella prepared the stimuli and helped in conducting the experiments; Phipps Arabie read the manuscript and gave many helpful suggestions.
Received for publication August, 1970.

1. S. S. Stevens, 1969, personal communication.

References

Cameron, A. T. 1947. The taste sense and the relative sweetness of sugar and other sweet substances. Sugar Research Foundation (New York), report 9.

Chappell, G. 1953. Flavour assessment of sugar solutions. *Journal of Science and Agriculture* 4:346–350.

Ekman, G., and Akesson, C. 1964. Saltness, sweetness, and preference: A study of quantitative relations in individual subjects. University of Stockholm Psychophysical Laboratory, report 177.

Engel, R. 1928. Experimentelle Untersuchungen über die Abhängigkeit der Lust and Unlust von der Reizstarke beim Geschmacksinn. *Archiv für die gesamte Psychologie* 64:1–36.

Feallock, J. B. 1965. Estimated magnitudes of taste under improved conditions of stimulus control. Unpublished doctoral dissertation, University of Virginia.

Gregson, R. A. M., and Russell, P. N. 1965. Psychophysical power law exponent for sucrose intensity. *Perceptual and Motor Skills* 20:294.

Indow, T., and Stevens, S. S. 1966. Scaling saturation and hue. *Perception and Psychophysics* 1:253–272.

Kocher, E. C., and Fisher, G. L. 1969. Subjective intensity and taste preference. *Perceptual and Motor Skills* 28:735–740.

Meiselman, H. L. 1968. Adaptation and cross-adaptation of the four gustatory qualities. *Perception and Psychophysics* 4:368–372.

Meiselman, H. L. 1971. Effect of presentation procedure on taste intensity functions. *Perception and Psychophysics* 10:15–18.

Moskowitz, H. R. 1968. Scales of intensity for single and compound tastes. Unpublished doctoral dissertation, Harvard University.

Moskowitz, H. R. 1970a. Ratio scales of sugar sweetness. *Perception and Psychophysics* 7:315–320.

Moskowitz, H. R. 1970b. Sweetness and intensity of artificial sweeteners. *Perception and Psychophysics* 8:40–42.

Panek, D. W., and Stevens, J. C. 1965. PSYCHOFIT, a computer program for the treatment of psychophysical data. Harvard University Laboratory of Psychophysics, report PPR 315. Program modified by Phipps Arabie.

Stevens, S. S. 1957. On the psychophysical law. *Psychological Review* 64: 153–181.

Stevens, S. S. 1969a. On predicting exponents for cross-modality matches. *Perception and Psychophysics* 6:251–256.

Stevens, S. S. 1969b. Sensory scales of taste intensity. *Perception and Psychophysics* 6:302–308.

Stevens, S. S., and Greenbaum, H. 1966. Regression effect in psychophysical judgment. *Perception and Psychophysics* 1:439–445.

Stone, H., and Oliver, S. 1969. Measurement of the relative sweetness of selected sweeteners and sweetener mixtures. *Journal of Food Science* 34:215–222.

continued from page 376

Prather, Hugh. *Notes to Myself*. Lafayette, Calif.: Real People Press, 1970. $2.00.

Raknes, Ola. *Wilhelm Reich and Orgonomy*. New York: St. Martin's Press, 1970. Pp. 130. $6.50.

Reger, Roger, ed. *Preschool Programming of Children with Disabilities*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xii, 123. $7.50.

Robinson, Floyd G. *Study Guide for Ausubel/Robinson School Learning*. New York: Holt, Rinehart and Winston, 1970. Pp. viii, 205.

Robinson, Francis P. *Effective Study*. 4th ed. New York: Harper & Row, 1970. Pp. xii, 303. $5.95.

Roche, Louis. *Le Traitement dans le service penal: Perspectives nouvelles*. Actes du X Congres francais de Criminologie, Lyon, 25–27 September 1969. Paris: Masson, 1970.

Rogers, Carl. *Carl Rogers on Encounter Groups*. New York: Harper & Row, 1970. Pp. vii, 172. $5.95; paperback, $1.95.

Roslansky, John D., ed. *Creativity*. Nobel Conference Lectures. New York: Fleet Academic Editions, 1971. Pp. 102. $7.50.

Rothman, Ester P. *The Angel Inside Went Sour*. New York: David McKay, 1971. Pp. 333. $7.95.

Rowland, Thomas G., and McGuire, J. Carson, eds. *The Mind of Man: Some Views and a Theory of Cognitive Development*. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xix, 167. $5.95.

Royce, J. R. *Toward Unification in Psychology*. Toronto: University of Toronto Press, 1970. Pp. 308. $12.50.

Rudensky, Morris, and Riley, Don. *The Gonif: Red Rudensky*. Blue Earth, Minn.: Piper, 1970. Pp. 215. $7.95.

Russell, Elbert W.; Neuringer, Charles; and Goldstein, Gerald. *Assessment of Brain Damage: A Neuropsychological Key Approach*. New York: Wiley, 1970. Pp. x, 167. $12.95.

Ryan, Thomas Arthur. *Intentional Behavior: An Approach to Human Motivation*. New York: Ronald Press, 1970. Pp. ix, 590.

Savage, E. Wade. *The Measurement of Sensation: A Critique of Perceptual Psychophysics*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 578. $15.00.

Schioldborg, Per. *Eye Movement in Visual Matching. A Temporal Pattern*. Oslo: Universitetsforlaget, 1969. Pp. 124. $6.00.

Schmitt, T. O.; Melnechuk, T.; Quarton, G. C.; and Adelman, G.; eds. *Neurosciences Research Symposium Summaries*. Vol. 4. Boston: The M.I.T. Press, 1970. Pp. 511. $12.50.

Scott, Edward M. *Struggles in an Alcoholic Family*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xiii, 265. $15.00.

Seidenberg, Robert. *Marriage in Life and Literature*. New York: Philosophical Library, 1970. Pp. ix, 307. $5.95.

Shave, David W. *The Language of Transference*. Boston: Little, Brown and Co. Pp. vii, 214.

Sherwood, Nancy M., and Timiras, Paola S. *A Stereotaxic Atlas of the Developing Rat Brain*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 209. $20.00.

completed on page 420

American Journal of Psychology
1971, Vol. 84, No. 3

ANCHORING EFFECTS IN PSYCHOLOGICAL SCALING

H. J. A. Rimoldi, M. B. Donnelly, and W. J. Laack
Loyola University, Chicago

When 122 subjects ranked 15 value statements in all 105 possible pairs from 0 to 100, as compared with when they ranked the same 15 statements from −4 to +4 by the successive-intervals method, a significant anchoring effect was found for only 7 of the 15 stimuli. In all 7 cases, the slope was negative (negative contrast effect) and the stimuli of low affective value.

Anchoring and contrast effects are rather old and well-established concepts (for a summary, see Guilford, 1954). One example is the work by Volkmann (1936), who found that when subjects rated the degree of inclination of a line of light, there was a shift in the scale values toward the anchoring stimulus (either real or imagined) upon its introduction. Cohen (1937) had subjects rate the pleasantness of odors on a seven-step category scale. In the first part of his experiment, 15 stimuli were rated. In the second part of the experiment, 6 unpleasant odors were added. These 21 stimuli were then rated. It was found that there was an upward shift in the values assigned to the 15 stimuli used in the first part. Postman and Miller (1945) obtained similar anchoring effects when they had subjects make temporal judgments. Their major findings indicated that subjects need not be informed of the presence of anchoring stimuli and that the stimuli close to the anchoring stimuli in affective value are more affected than are the remote ones.

More recently, Heerman, Bartlett, and Rettig (1962) have demonstrated anchoring effects in the scaling of occupations. Using an analysis-of-variance approach, they found that the scale values assigned to five stimuli of moderate value were lowered in the presence of five high anchors and raised when the subjects were presented with five lower anchors. In a different context, Hicks and Campbell (1965) found that estimates of zero points were markedly affected by anchoring stimuli when subjects rated (*a*) desirability of birthday gifts, (*b*) traffic violations, and (*c*) behavioral symptoms. However, Jones (1967) found that these effects could be greatly reduced by using a log-normal model (Case VI).

While these results are rather clear-cut and consistent, the problem of interpretation remains. That is, do these results indicate some psychological phenomenon? Or are they simply results of the scaling procedures, as has been suggested by several of the authors cited above?

The typical approach in these experiments has consisted of having the subjects rate n stimuli in m categories, adding to these a new set of n_1 stimuli, and then having the same subjects rate the $n + n_1$ stimuli in the same m categories. However, the change in scale values obtained under this procedure might be an artifact independent of the true affective value of the stimuli. More simply, the n stimuli might just have had to 'make room' for the n_1 new stimuli, with a resultant shift of the original scale values.

The purpose of the present study was to examine whether stimulus i influences the affective value assigned to stimulus j when subjects are aware of the entire range of stimuli to be rated. The anchoring effect was analyzed using the concept of independence. That is, if the scale value of a stimulus does not change when presented with another stimulus, then the stimuli may be considered independent. In this case no anchoring effect will be observed. If, on the other hand, a stimulus does change upon the introduction of some other stimulus, we may say that the affective value of the original stimulus is *dependent* upon the added stimulus. In this case the anchoring effect will appear. This work began in 1955 and several articles have already been presented (Devane and Rimoldi, 1960; Donnelly and Rimoldi, 1967). In fact, a preliminary report of this study appeared some time ago (Devane and Rimoldi, 1961).

METHOD

—Subjects—The subjects were 122 men and women, college students from several introductory psychology courses.

—Procedure—Fifteen value statements were prepared using the Spranger and the Allport-Vernon descriptions of ideal types of value systems.[1] These value statements were presented to each subject in a booklet including four sections. The first section, *part A*, consisted of 105 pairs made by combining the 15 single statements in all possible pairs. Each subject was instructed to assign any value between 0 and 100 to *each* member of the pair in accordance with his degree of agreement with the statement. The subjects were specifically instructed to assign these values only after reading carefully each statement in the pair. They were told not to use the extreme values unless they felt complete agreement or disagreement with the statement. Part B consisted of the same 105 pairs, but the subjects were instructed to assign to each statement in any pair any value between 0 and 100, provided that the sum of the values added up to 100. This is the constant-sum method. Other instructions were similar to those in Part A. In Part C, the subjects were asked to give a value from 0 to 100 to each of the pairs as such. The last section, *part D*, consisted of the 15 single statements, which the subjects were to rate using the successive-intervals method (Rimoldi and Hormaeche, 1955). The range of categories was from −4 to +4.

Parts A, B, C and D were administered in the same experimental session. The order of presentation of the 15 single statements (part D) and of the 105 pairs of statements (parts A, B and C) was random but identical for all subjects. The results obtained in parts B and C have been reported elsewhere (Donnelly and Rimoldi, 1967) and are not repeated here.

—Design of the study—Let $X_{p(j)}$ be the score assigned by subject p to stimulus j, where $p = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, n$. Let $X_{p(j/i)}$, $j \neq i$, be the value assigned by subject p to stimulus j in the presence of stimulus i. If $X_{p(j/i)} = X_{p(j)}$, we may conclude that j and i are independent, in the sense that the presence of i does not affect the value of j.

In part A there were $n(n-1)/2$ pairs of stimuli. For every (j/i) combination we obtain

$$\frac{\sum_{p=1}^{N} X_{p(j/i)}}{N} = M_{(j/i)}, \quad (j \neq i) . \qquad [1]$$

The notation $M_{(j/i)}$ refers to the average value of stimulus j when judged in the presence of stimulus i. In general, the i subscript identifies the *condition* under which stimulus j is judged, where the condition is defined in terms of the affective value of stimulus i. Thus each one of the 15 statements functioned both as a stimulus in a particular condition and as a condition for the other 14 stimuli. For example, in the pair in which stimulus 5 and stimulus 2 were rated, two $M_{(j/i)}$ values, $M_{(5/2)}$ and $M_{(2/5)}$, were obtained. That is, we obtained the value of stimulus 5 when paired with stimulus 2 and the value of stimulus 2 when paired with 5.

Since each $M_{(j/i)}$ value is based on the same N, averaging the $M_{(j/i)}$ values for all the i's, we obtain

$$\frac{\sum_{i=1}^{n} M_{(j/i)}}{N-1} = M_j, \quad (i \neq j). \qquad [2]$$

Each of the 15 stimuli was also judged using the successive-intervals approach of part D.

The entries in the cells of Table 1 correspond to the $M_{(j/i)}$ values as defined in Equation 1 and obtained from part A. Table 1 is not symmetrical, since $M_{(j/i)}$ may be different from $M_{(i/j)}$. The entries in the row vectors, α_j, describe the mean values assigned to stimulus j under each of the conditions, so that $\alpha_j = (M_{(j/1)}, M_{(j/2)}, \ldots, M_{(j/n)})$. Systematic changes in the entries of any of these α_j vectors may then be referred to as demonstration of an anchoring effect. The effect is defined in terms of changes in $M_{(j/i)}$ as i varies in affective value.

Table 1. Data model

	Condition							
Stimuli	1	2	3 . . .	i	. . . n	θ	γ	δ
α_1	$M_{1/1}$	$M_{1/2}$	$M_{1/3}$	$M_{1/i}$	$M_{1/n}$	M_1	${}^{\sigma}\alpha_1$	δ_1
α_2	$M_{2/1}$	$M_{2/2}$	$M_{2/3}$	$M_{2/i}$	$M_{2/n}$	M_2	${}^{\sigma}\alpha_2$	δ_2
α_3	$M_{3/1}$	$M_{3/2}$	$M_{3/3}$	$M_{3/i}$	$M_{3/n}$	M_3	${}^{\sigma}\alpha_3$	δ_3
$\vdots$								
α_j	$M_{j/1}$	$M_{j/2}$	$M_{j/3}$	$M_{j/i}$	$M_{j/n}$	M_j	${}^{\sigma}\alpha_j$	δ_j
$\vdots$								
α_k	$M_{k/1}$	$M_{k/2}$	$M_{k/3}$	$M_{k/i}$	$M_{k/n}$	M_k	${}^{\sigma}\alpha_k$	δ_k
$\vdots$								
α_n	$M_{n/1}$	$M_{n/2}$	$M_{n/3}$	$M_{n/i}$	$M_{n/n}$	M_n	${}^{\sigma}\alpha_n$	δ_n
δ'	δ'_1	δ'_2	δ'_3	δ'_i	δ'_n			

The entries in column θ of Table 1 define a vector of M_j values as defined in Equation 2. These M_j values are used as estimates of the diagonal entries in Table 1; that is, $M_{(j/i)}$ where $j = i$. The entries in column γ correspond to the standard deviation of the entries in each α vector. They are indices of changes in the value of a fixed stimulus j under all conditions.

The M_j values are not experimentally independent of the $M_{(j/i)}$ values. In column δ and row δ' the entries correspond to an independent estimate of the scale value for each of the n stimuli. These were obtained from part D. These values are designated by either δ_j $(j = 1, 2, \ldots, n)$ or δ'_i $(i = 1, 2, \ldots, n)$; δ_j is used whenever we are referring to an independent estimate of the affective value of the stimulus being judged, while δ'_i is employed to define the condition under which stimulus j is presented. That is, each condition is defined in terms of an independent estimate of the affective value of stimulus i.

If all the entries of vector α_j are identical, then the effect of the i conditions on stimulus j is null. If this is not the case, it may be hypothesized that the value given to stimulus j varies with the i conditions. The question is, What is the functional correspondence between the affective value of the various i conditions and the scale values for each fixed j? This correspondence can be examined by plotting each of the α vectors $M_{(j/i)}$ for a fixed α_j against row δ'.

Assuming a linear model, the relationship between α_j, for a fixed j, and δ', can be stated as

$$M_{(j/i)} = b\delta'_i + c_j , \tag*{[3]}$$

where c_j and b are constant such that

$$c_j = a\delta_j + k , \tag*{[4]}$$

where a and k are constant.

If $b = 0$, then the value of stimulus j in the presence of i is the same for all i conditions; that is, $M_{(j/i)} = a\delta_j + k$ for all i's. If b is positive, then the affective value of $M_{(j/i)}$ increases as the value of i increases (positive anchoring effect). Finally, if b is negative, the value of $M_{(j/i)}$ decreases as i increases (negative anchoring effect).

It is of interest to know which j stimuli are similarly affected by the different i conditions. For instance, let the stimuli be j and k. Then the correlation between the $M_{(j/i)}$ and $M_{(k/i)}$ values for varying i's provides an indication of the relationship between j and k. Estimations of these correlations were obtained for all the $n(n-1)/2$ combinations of the α vectors.

RESULTS

Table 2 presents the experimental results according to the model set forth in Table 1. The relationship between the category means M_j, from column θ of Table 2, and the δ_j values obtained from part D is given in Figure 1, together with the pertinent statistics.

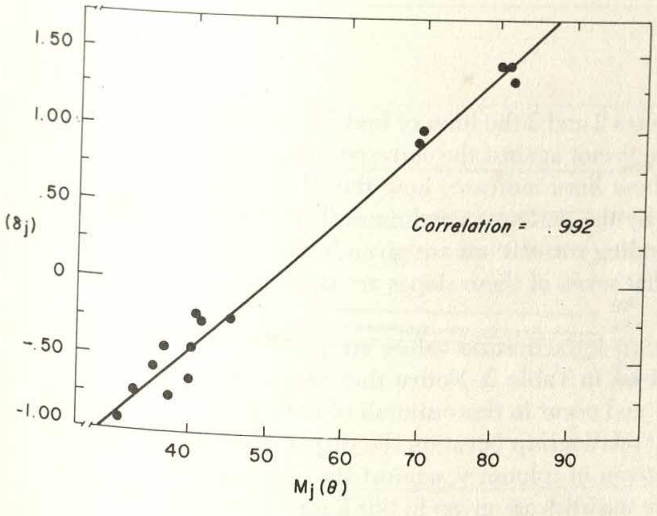


Fig. 1. M_j versus δ_j

Table 2. Data summary

	Condition								
Stimuli	1	2	3	4	5	6	7	8	9
α_1	(31.61)	35.00	30.00	33.28	31.15	34.92	32.38	35.41	31.88
α_2	35.49	(33.06)	32.46	32.54	34.92	33.77	34.18	34.18	35.25
α_3	39.92	42.21	(38.82)	42.21	41.07	37.38	42.54	28.44	37.21
α_4	45.57	42.29	39.43	(41.83)	40.41	42.95	44.34	42.79	42.05
α_5	32.91	37.29	34.26	38.20	(36.37)	37.95	37.21	36.80	38.77
α_6	41.07	39.59	38.03	37.13	37.54	(37.63)	40.98	35.94	36.97
α_7	42.54	41.23	44.84	39.84	41.48	45.98	(41.77)	41.97	42.21
α_8	45.33	43.69	44.75	40.08	44.92	42.13	45.00	(42.52)	42.13
α_9	45.90	47.87	47.87	46.72	46.48	48.93	45.16	45.90	(46.26)
α_{10}	42.05	42.70	43.36	43.61	41.15	43.03	44.18	43.03	43.77
α_{11}	71.97	71.23	69.43	68.36	68.36	70.25	69.43	68.36	67.29
α_{12}	71.48	67.87	68.11	69.26	68.03	70.93	70.25	70.66	69.03
α_{13}	82.70	82.79	84.75	82.29	87.21	82.95	80.41	81.97	81.89
α_{14}	83.69	81.97	81.48	81.97	82.38	83.44	81.48	80.57	82.29
α_{15}	83.20	82.38	83.36	82.21	82.38	82.95	82.29	82.54	81.56
δ'_i	−.934	−.796	−.751	−.650	−.588	−.478	−.454	−.290	−.278

* $p < .05$.

In Figures 2 and 3 the lines of best fit obtained by plotting the entries of each α_j vector against the corresponding entries of δ' are given. Each one of these lines indicates how the $M_{(j/i)}$ values (for a fixed j) are affected by the various i conditions. The slopes of these lines and the corresponding correlations are given in the last two columns of Table 2. Notice that seven of these slopes are significantly different from a slope of zero.

In Figure 4, the δ_j scale values are plotted against the corresponding slopes given in Table 2. Notice that the slopes that are significant are negative and occur in those stimuli of low affective value. Figure 5 presents the relationship between the dispersion within each of the α vectors, as given in column γ, against the corresponding δ_j value. The appropriate statistics are given in this figure. Note that there is a tendency for those stimuli of low affective value to have a greater variability in their $M_{(j/i)}$ values.

11	12	13	14	15	$\theta(M_j)$	$\gamma(\sigma_j)$	δ	Slopes	Correlations
.00	30.33	28.85	26.31	28.69	31.61	2.66	−.934	−.256*	−.77
.79	30.00	29.92	33.36	30.49	33.06	1.77	−.796	−.134*	−.71
.79	40.33	33.61	37.46	40.82	38.82	3.83	−.751	−.119	−.25
.84	41.56	40.25	39.51	40.98	41.83	1.82	−.650	−.124*	−.56
.29	38.77	31.48	35.00	35.90	36.37	2.13	−.588	−.053	−.12
.30	37.46	34.26	34.84	34.18	37.63	2.09	−.478	−.203*	−.80
.64	39.51	40.90	39.10	39.43	41.77	2.00	−.454	−.142*	−.59
.18	38.77	41.06	41.48	39.34	42.52	2.11	−.290	−.155*	−.61
.54	45.90	44.67	43.36	44.02	46.24	1.70	−.278	−.098	−.48
.74	41.72	39.51	42.05	41.56	42.32	1.26	−.274	−.096*	−.63
.58)	69.59	71.31	71.31	69.43	69.58	1.37	.788	.034	.20
.59	(69.89)	70.33	73.28	69.67	69.89	1.41	.848	.075	.43
.79	80.82	(82.79)	82.79	82.70	82.79	1.58	1.230	−.050	−.24
.29	81.72	82.79	(82.18)	81.97	82.18	.77	1.310	−.006	−.06
.05	81.48	82.62	81.88	(82.50)	82.50	.57	1.318	−.023	−.30
.788	.848	1.230	1.310	1.318					

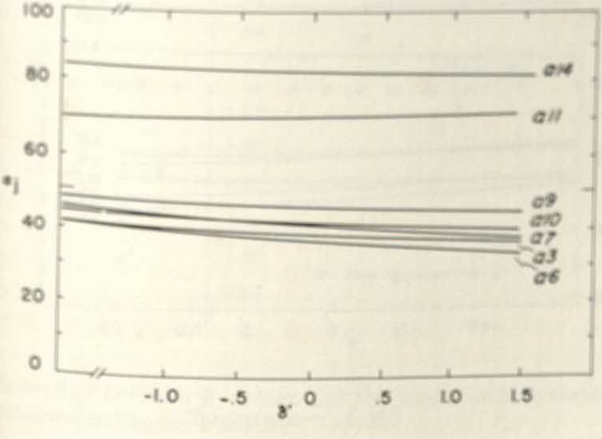


Fig. 2. α_j versus δ'

Table 3. Intercorrelations of α vectors

α	$\alpha6$†	$\alpha1$†	$\alpha10$†	$\alpha8$†	$\alpha2$†	$\alpha7$†
$\alpha6$†						
$\alpha1$†	.753*					
$\alpha10$†	.532*	.532*				
$\alpha8$†	.675*	.233	.169			
$\alpha2$†	.631*	.470	.494	.710*		
$\alpha7$†	.555*	.485	.346	.539*	.392	
$\alpha4$†	.753*	.776*	.460	.261	.483	.347
$\alpha9$	.394	.461	.167	.380	.208	.565*
$\alpha15$	.324	.241	−.195	.615*	.112	.709*
$\alpha3$	.431	.008	.204	.157	.015	−.097
$\alpha13$	−.172	−.123	−.388	.391	.237	.276
$\alpha5$	.197	.536*	.543*	−.249	.171	.008
$\alpha14$	−.038	−.088	−.338	.146	.171	.310
$\alpha11$	.020	−.354	−.419	.139	−.197	−.116
$\alpha12$	−.147	−.314	−.100	−.214	.062	−.100

*$p > .05$. † Significant contrast effect for this variable.

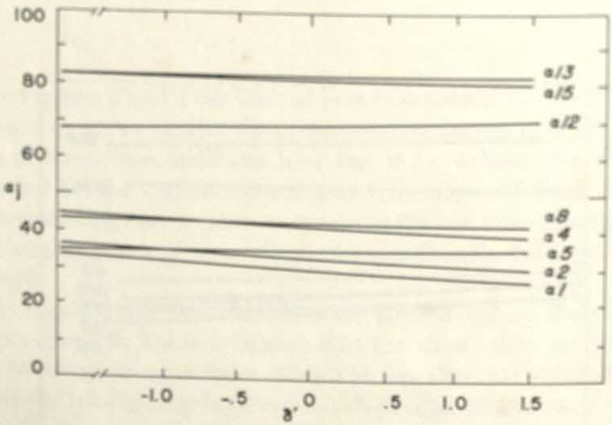


Fig. 3. — α_j versus δ'

α9	α15	α3	α13	α5	α14	α11
.213						
.036	−.021					
.208	.373	.108				
.352	−.450	.288	−.517			
.148	.057	.235	.216	−.308		
−.072	.161	.000	−.078	−.632*	.472	
−.491	−.073	−.352	−.410	−.299	.353	.441

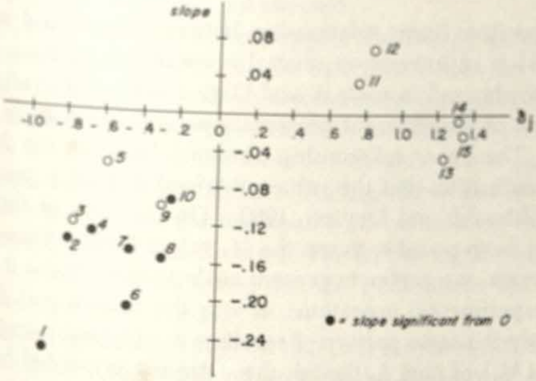


Fig. 4. Relationship between δ_j values and the corresponding slopes (α_j versus δ') indicated in Figures 2 and 3

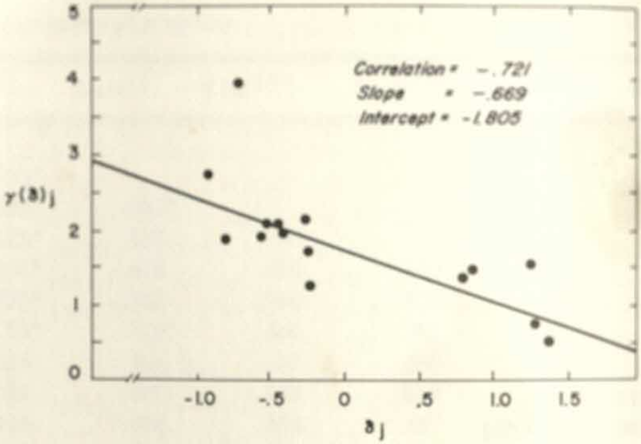


Fig. 5. Relationship between δ_j and $\gamma(\sigma_j)$

The correlations between the entries in the α vectors are presented in Table 3, where the α vectors have been ordered in terms of the significance of their slopes. Notice that the variables that have lower scale δ_j values tend to be significantly correlated among themselves.

DISCUSSION

The high positive linear relationship between the δ_j and M_j values (see Figure 1) is of interest on several accounts. It indicates that the different methods used in parts A and D give very closely related estimates, in spite of the different presentations and instructions given to the subjects. The linear relationship obtained between the M_j and δ_j values is an indication that the values obtained in part A are normally distributed (Rimoldi and Devane, 1960). On the basis of this finding, it would have been possible to use the M_j values for the discussion that follows; however, we prefer to present scale values obtained using independent experimental operations, as was the case in part D. Nevertheless, exactly the same pattern of results was obtained using the category means (M_j) of part A, though these are not presented here.

The lines of best fit obtained by plotting the entries in the rows of Table 2 against the corresponding δ'_i values are approximately linear, except for α_3. Some of the slopes are positive, though not significantly different from zero. The negative slopes for stimuli 1, 2, 4, 6, 7, 8, and 10 are significantly different from zero beyond the .05 level. The slope in these graphs may be interpreted as the change that occurs in the $M_{(j/i)}$ values when subjects are judging stimulus j in the presence of the various i's. A negative slope can be interpreted as indicating that the values assigned to j under the various i conditions decrease as the scale value corresponding to condition i increases. That is, if condition i has a high scale value, the value given to j is lower than when i has a low value. This inverse relationship is referred to as a negative contrast effect. A slope of zero indicates that the value of j is unaffected by any i condition, while a positive slope indicates that the value assigned to j increases as the scale value of condition i increases.

By these criteria, a significant effect was found for only 7 of the 15 stimuli. In all these cases, the slope was negative (negative contrast effect). The stimuli for which a contrast effect was found were all stimuli of low affective value. Stimuli of higher affective value were relatively unaffected by the i conditions. Inspection of Figure 4 seems to show a positive relationship between the value of the slope and the scale value of the stimuli. Furthermore, a biserial correlation of −.88 was obtained between the δ_j value and the presence or absence of a significant contrast effect.

Figure 5 seems to show that the values δ_j are negatively related to the dispersion ($\sigma_{\alpha j}$) of the $M_{(j/i)}$ values. This result can be interpreted in terms of the contrast effect. If a stimulus is affected differently by the i conditions, then we would expect a greater dispersion in the $M_{(j/i)}$ values than in those stimuli that are not affected.

The results in Table 3 further support the above results. Notice that the sub-matrix corresponding to the intercorrelations between those α vectors in which a contrast effect was detected contains all positive correlations, most of which are significant beyond the .05 level. The remaining entries in this table do not seem to follow any clear pattern, nor do many of them approach significance. The significant positive correlations indicate that the i conditions had similar effects on all of the j stimuli in which a contrast effect was detected. The lack of significant correlations in the rest of the table indicates that the i conditions did not produce similar effects on these stimuli.

In summary, our findings suggest that only the stimuli at the lower end of the scale were significantly affected by the presence of other stimuli. Specifically, their affective value was decreased when they were presented with stimuli of higher value. This is further verified by the fact that stimuli of low scale values had greater dispersions over the i conditions than did stimuli with high scale values. Nevertheless, it does not follow that stimuli with high dispersions necessarily show a linear contrast effect, as was demonstrated by stimuli 3, 5, and 9. In the case of stimulus 3, it was found that by plotting the $M_{(3/i)}$ values against the δ'_i values, a nonlinear relationship was obtained. It is interesting that statement 3 was the only disjunctive statement in the set—which fact may account for this result.

Thus, in terms of our original hypothesis, and for these data, the following conclusions may be posited: first, that if δ_j corresponds to a stimulus of low affective value, the probability is that b in Equation 3 will be negative; and second, that if δ_j corresponds to a stimulus of moderate or high affective value, the probability is that b will tend to zero. That is, the affective value of moderate and high stimuli is independent of the i conditions, whereas this is not true of stimuli of lower affective value.

Unfortunately, the δ_j values (or M_j values) given in Table 2 show that the original stimuli tended to define two clearly separated groups: one with low scale values and the other with high scale values, so that the middle range was not clearly represented. It will be appropriate to investigate contrast effects by using the concept of independence and a set of stimuli more evenly distributed throughout the whole range.

It should be noted that we could have interpreted the data in terms of a fixed i and varying αs. In typical experiments in this area, this is the usual way to analyze the data. In fact, the data were analyzed from this point of view; these analyses were not presented because they indicate exactly the same results as those we have just reported. However, the interpretation becomes more complex.

These results may have application in the areas of public-opinion polling, advertising, market research, politics, and so on. That is, if we can generalize from our results, it may not be the most effective policy to present a person or object of low affective value with one of high value if we are trying to 'sell' the one of low value. Also, an object or person with a 'high value' may not be enhanced by juxtaposition with objects or persons of low or equal value.

Notes

Dr. Rimoldi is now the Director of the Centro Interdisciplinario de Investigaciones Científicas y Técnicas (CIIPME) of the National Research Coun Argentina and the School of Sciences, University of Buenos Aires; and D. Donnelly is now at Wayne State University. Received for publication March 9, 1970; revised.

References

Devane, J. R., and Rimoldi, H. J. A. 1960. The determination of interval limits. Loyola University Psychometric Laboratory, publication 16.

Devane, J. R., and Rimoldi, H. J. A. 1961. The experimental study of value stimuli in combination. Loyola University Psychometric Laboratory, publication 20.

Donnelly, M. B., and Rimoldi, H. J. A. 1967. Experimental considerations concerning category and magnitude scaling. *American Journal of Psychology* 80:602–607.

Cohen, N. E. 1937. The relativity of absolute judgments. *American Journal of Psychology* 49:93–100.

Guilford, J. P. 1954. *Psychometric methods.* New York: McGraw-Hill.

Heerman, E. F., Bartlett, C. J., and Rettig, S. 1962. The applicability of anchorage effects to the scaling of occupations. *Journal of Social Psychology* 56:115–120.

Hicks, J. M., and Campbell, D. 1965. Zero-point scaling as affected by social object, scaling method, and context. *Journal of Personality and Social Psychology* 2:793–808.

Jones, L. V. 1967. Invariance of zero-point scaling over changes in stimulus context. *Psychological Bulletin* 67:153–164.

Postman, L., and Miller, G. A. 1945. Anchoring of temporal judgments. *American Journal of Psychology* 58:43–53.

Rimoldi, H. J. A., and Devane, J. R. 1960. Some considerations on scaling procedures. *Perceptual and Motor Skills* 11:207–213.

Rimoldi, H. J. A., and Hormaeche, M. 1955. The law of comparative judgment in the successive intervals and graphic rating scale methods. *Psychometrika* 20(4):307–318.

Volkmann, J. 1936. The anchoring of absolute scales. *Psychological Bulletin* 33:742–743.

continued from page 406

Shiloh, Ailon, ed. *Studies in Human Sexual Behavior: The American Scene.* Springfield, Ill.: Charles C Thomas, 1970. Pp. xxviii, 460. $16.50.

Shull, Fremon A., Jr.; Delbecq, Andre L.; and Cummings, L. L. *Organizational Decision Making.* New York: McGraw-Hill, 1970. Pp. xvi, 320. $11.95.

Small, Leonard. *The Briefer Psychotherapies.* New York: Brunner/Mazel, 1971. Pp. xx, 262. $7.50.

Smith, Kendon. *Behavior and Conscious Experience.* Athens: Ohio University Press, 1969. Pp. 139. $6.00.

Stevens, Barry. *Don't Push the River.* Lafayette, Calif.: Real People Press, 1970. Pp. 268. $7.00; paperback, $3.50.

Strackbein, O. R. *Permissiveness.* New York: Vantage Press, 1970. Pp. 185. $4.95.

Straus, Erwin W., and Griffith, Richard M., eds. *Phenomenology of Memory: The Third Lexington Conference on Pure and Applied Phenomenology.* Pittsburgh: Duquesne University Press, 1970. Pp. xiv, 205. $6.95.

Sutton-Smith, Brian, and Rosenberg, B. G. *The Siblings.* New York: Holt, Rinehart and Winston, 1970. Pp. ix, 198. $7.50.

Swingle, Paul, ed. *The Structure of Conflict.* New York: Academic Press, 1970. Pp. x, 305. $12.50.

Voeks, Virginia. *On Becoming an Educated Person: The University and College.* 3rd ed. Philadelphia: W. B. Saunders, 1970. Pp. 278. $2.95.

Wertheimer, Michael. *A Brief History of Psychology.* New York: Holt, Rinehart and Winston, 1970. Pp. ix, 163. Paperback, $2.50.

White, Burton L. *Human Infants: Experience and Psychological Development.* Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xv, 160. $4.95.

Whittaker, James O. *Student's Workbook to Accompany "Introduction to Psychology.* 2d ed. Philadelphia: W. B. Saunders, 1970. Pp. 211. $3.75.

Williams, Moyra. *Brain Damage and the Mind.* Baltimore: Penguin Books, 1970. Pp. 175. Paperback, $1.95.

Wirt, Frederick M. *Politics of Southern Equality: Law and Social Change in a Mississippi County.* Chicago: Aldine, 1970. Pp. x, 335. $10.00.

Wiseman, Jacqueline P. *Stations of the Lost: The Treatment of Skid Row Alcoholics.* Englewood Cliffs, N.J.: Prentice-Hall, 1970. Pp. xiv, 345. $8.95; paperback, $4.95.

Wolff, Kurt. *The Emotional Rehabilitation of the Geriatric Patient.* Springfield, Ill.: Charles C Thomas, 1970. Pp. xiv, 232. $11.50.

Wolff, Kurt, ed. *Patterns of Self-destruction: Depression and Suicide.* Springfield, Ill.: Charles C Thomas, 1970. Pp. xi, 88. $6.50.

Wortis, Joseph, ed. *Mental Retardation: An Annual Review.* New York: Grune & Stratton, 1970. Pp. xii, 228. $14.75.

Wright, D. S.; Taylor, Ann; Davies, D. Roy; Sluckin, W.; Lee, S. G. M.; and Reason, J. T. *Introducing Psychology: An Experimental Approach.* New York: Penguin Books, 1970. Pp. 736. Paperback, $3.95.

Young, Geryke. *Two Worlds—Not One: Race and Civilization.* London: Ad Hoc Publications, 1969. Pp. 144. $2.00.

American Journal of Psychology
1971, Vol. 84, No. 3

A METHOD TO ASSESS CHANCE ALTERNATION LEVELS IN THE FORCED-TRIAL PARADIGM

Susan Karp Manning
Hunter College of the City University of New York

Improper estimation of chance alternation levels in the forced-trial paradigm has undoubtedly led to incorrect conclusions about the amount of spontaneous alternation present in this paradigm as compared to chance and as compared to the free-trial paradigm. A method that may be used to obtain the chance level of alternation in the forced-trial paradigm is proposed.

Consider an experimental paradigm consisting of one or more widely spaced trials. Each trial is made up of two successive runs in a two-choice situation, such as a T-maze. It has been found that animals, whether rewarded or not, tend on the second run to choose the alternative opposite from their first choice (Dember and Fowler, 1958). The excess of alternate responses over the chance level has been termed spontaneous alternation. This phenomenon has been studied in two types of situations, referred to as the free-trial and forced-trial paradigms respectively. In the *free-trial* paradigm, the animal gets two free runs. In the *forced-trial* situation, on the first run the animal is required to make a preselected response (e.g., one goal arm may be blocked off, forcing the animal to the other arm), while on the second run the animal is given a free choice.

A good deal of research has been done using the two paradigms. Unfortunately, the forced-trial paradigm has generally been thought to have the same chance level of alternation as the free-trial paradigm. This has led such researchers as Hughes (1966) and Thompson (1962) to make direct—and invalid—comparisons between alternation levels in the two paradigms. Even Dember and Fowler (1959), who were concerned with appropriate measures of chance alternation in the free-trial case, compared the chance levels of alternation in the free- and forced-trial paradigms as though they were identical. Further, there has never been a reasonable attempt to conceptualize the forced-trial paradigm or to develop a general procedure for obtaining alternation in the forced-trial case.

To begin with, most early studies and some current ones (Still and MacMillan, 1969) assumed the chance level of alternation to be .5 for the free-trial paradigm. Much research has shown, however, that in a two-choice situation, such as a T-maze, many factors can cause animals to choose one side consistently more often than the other. Animals reliably choose black stimuli more often than white ones, frequently develop so-called position preferences, and may even show preferences for such stimuli as the odor of peppermint over that of pipe tobacco (Douglas and Isaacson, 1965).

Since the opportunity for the development of choice inequalities is common in experimental situations (due to the use of differentiated goal arms or to other factors inherent in experimental settings), Dember and Fowler (1958) pointed out that the assumption of a .5 chance level of alternation is frequently inaccurate for the free-trial paradigm. They reasoned that an animal with a .8 right position habit would be expected to go to the right arm of a maze .8 of the time on each of the two free trials if the trials were independent (i.e., if the performance was chance). Thus the probability of alternation for such an animal would be the chance probability of a right response followed by a left response. In this case, that would be $(.8 \times .2) + (.2 \times .8)$, or .32.

Using a more precise notation, let us arbitrarily designate the two choices available to the animal as M and O. Assume the probability of choice of alternative M to be m, of alternative O to be $1 - m$. Thus the general formula for expected alternation in the free-trial case is $m(1 - m) + (1 - m)m$.

In their use of this formula, Dember and Fowler, and others following their lead, assessed the probability of choice of M (and O) by taking the proportion of choices on free first runs over the total number of choices the animal makes. Since a sample proportion is an unbiased estimate of a population proportion (Sanders, 1958), this may be assumed to be a reasonably valid measure of m, depending, of course, on the number of trials used for estimation for its accuracy. Further complications in assessing chance alternation in the free-trial paradigm have been pointed out by Koppenaal (1962); they are not discussed here because they are not related to the forced-trial situation.

This analysis of the free-trial paradigm has elements of importance for the forced-trial paradigm, but it does not correctly analyze, or provide the correct method for assessing, chance levels of alternation in that forced-trial situation. It is true that, as in the free-trial paradigm, the chance level of alternation in the forced-trial situation depends in part on the previously defined quantity m. However, still another parameter is relevant here.

Let us define f as the proportion of runs that an animal is forced toward m. Accordingly, forcing toward O occurs on a proportion of runs equal to $1 - f$. Note that f is a constant, predetermined by the experimenter. Thus it is known, not estimated. For example, if the experimenter forces the animal to the black (M) side of a T-maze three-fourths of the time, then $f = .75$. Since the probability of alternation in the forced-trial paradigm, as in the free-trial paradigm, is equal to the probability of a response to M being followed by a response to O plus the converse, the chance alternation in the forced-trial paradigm is thus

$$\begin{aligned}&f(m)(1-m) + f(1-m)(1-m) + (1-f)(1-m)(m) + \\ &(1-f)(m)(m)\ , \\ &\text{or } f(1-m) + m(1-f)\ . \end{aligned} \tag{1}$$

Given Equation 1, we may see, for example, that an animal with a preference for M that differed greatly from .5 would have a chance level of alternation that is low if forcing were in the direction of the most frequent choice, and high if the reverse were true. In the extreme case, an animal that always chose M would never alternate if forced toward M, and would always alternate if forced toward O. Intermediate values of f and m produce intermediate results. An f of .5 always leads to .5 chance alternation. This and other situations may be easily derived by examining and working with Equation 1.

The difference between this method for the forced-trial paradigm and that for the free-trial paradigm may be illustrated with the following example. Assume a single animal with a known .7 preference for the black side of a T-maze (side M). If this animal is forced toward black half of the time and toward white half of the time, the chance level of alternation is .5 using Equation 1. Using Dember and Fowler's procedure and equation, the expected alternation for *free*-trial animals is .42. Thus, direct comparisons of relative alternation in the two paradigms may not be validly made. Many other examples are possible.

The analysis above is described in terms of a single animal. But unlike the equation for assessing alternation in the free-trial paradigm (Koppenaal, 1962), Equation 1 is also valid for predicting group alternation even when individual preferences for M vary. The value of m for each subject may be obtained, and a mean or group m then calculated. Provided that f is constant for all subjects, this mean value may be used to obtain the chance level of alternation for the group.

It is thus suggested that future research involving the forced-trial paradigm use Equation 1 to assess the chance level of alternation, especially when it is of interest to know whether alternation is 'significantly' different from chance. It will also be important to take into account the differing chance levels of alternation when the free- and forced-trial paradigms are to be compared.

It should be further noted that chance alternation in the forced-trial paradigm depends on m as well as f. Thus in all cases other than an equal proportion of forcing to both M and O ($f = .5$), it will be necessary to get estimates of m. This may involve giving the animal independent (spaced) free trials in addition to the forced ones.

Notes

Received for publication April 7, 1970; revised October 26, 1970.

References

Dember, W. N., and Fowler, H. 1958. Spontaneous alternation behavior. *Psychological Bulletin* 55:412–428.

Dember, W. N., and Fowler, H. 1959. Spontaneous alternation after free and forced trials. *Canadian Journal of Psychology* 13:151–154.

Douglas, R. J., and Isaacson, R. L. 1965. Homogeneity of single-trial response tendencies and spontaneous alternation in the T-maze. *Psychological Reports* 16:87–92.

Hughes, R. N. 1966. Stimulus-change effects of forcing in alternation behavior. *Psychological Reports* 19:515–518.

Koppenaal, R. J. 1962. On the determination of expected probabilities of alternation. *Psychological Reports* 11:666.

Senders, V. 1958. *Measurement and statistics.* New York: Oxford University Press.

Still, A. W., and MacMillan, A. 1969. Odor trial and spontaneous alternation. *Psychonomic Science* 16:160–161.

Thompson, M. E. 1962. Alternation in the T-maze as a function of three variables. *Psychological Reports* 7:103–110.

American Journal of Psychology
1971, Vol. 84, No. 3

ATTRIBUTE-RESPONSE CORRELATIONS IN CONCEPT ATTAINMENT

Richard L. Gottwald
Yale University

The attainment of three concepts was studied with all irrelevant attributes uncorrelated, or with up to two irrelevant attributes partly correlated, with the responses. The correlated irrelevant attributes increased trials to criterion or led to new solutions based on these attributes. All-or-none learning was evident when individual attributes (relevant or irrelevant) had zero or perfect correlations with the responses, but not when one or more attributes had intermediate correlations. Almost all the subjects' solutions used combinations of affirmations and conjunctions rather than disjunctions and other complex rules that would also have been correct.

The concept-attainment task requires that the subject learn to assign a set of responses to some larger set of stimuli. Usually the stimuli vary along several discrete attributes and the subject is expected to learn a rule involving one or more of these attributes to assign the responses to the appropriate stimuli. Most concept-attainment experiments involve two classes of attributes—relevant and irrelevant. The relevant attributes are those that are pertinent to the conceptual rule. The irrelevant attributes are those that are not pertinent to the conceptual rule, and usually they are orthogonal to both the responses and the relevant attributes.

It is not necessary, however, for irrelevant attributes to be orthogonal to the responses, and frequently in real life the irrelevant attributes are partly correlated with them. Several investigators have manipulated the magnitude of the correlation between the irrelevant attributes and the responses, but the results are inconsistent. Bourne and Haygood (1960) varied the correlation of an irrelevant attribute in a single-attribute two-response concept. They found that an increase in the correlation produced a decrease in both number of errors and trials to criterion. But Gormezano and Grant (1958) and Abraham, Gormezano, and Wiehe (1964) varied the correlation of an irrelevant attribute in the four-response Wisconsin Card Sorting Test and found that an increase in the correlation had no effect on number of errors but did increase trials to criterion.

The present study was designed to further investigate the changes in concept attainment produced by having irrelevant attributes correlated with the responses. Further, the studies described above used concepts involving a single relevant attribute, and in such concepts the relevant attribute is perfectly correlated with the responses. In certain other concepts, the relevant attributes are not perfectly correlated with the responses. Since it is thereby possible that the effects of the correlated irrelevant attributes are affected by the correlation of the *relevant* attributes, it was also decided to use several different concepts to achieve generality.

For an example of how the correlation between relevant attributes and the responses may vary across concepts, consider the set of four stimuli generated by combining two colors (red and green) with two shapes (triangle and square). Given this set, we observe that there are three meaningfully different rules (concepts) for assigning two responses, alpha and beta, to the stimuli; see Table 1. The simplest rule is termed an *affirmation*. By this rule, the set of stimuli associated with each response is specified by one value on the single relevant attribute (alpha is red). The relevant attribute in this concept is, of course, perfectly correlated with the responses.

Table 1. Three rules for assigning two responses to a stimulus set generated by two binary attributes

Attributes		Response assignments		
Color	Shape	Affirmation	Conjunction/disjunction	Biconditional
Red	Triangle	α	α	α
Red	Square	α	β	β
Green	Triangle	β	β	β
Green	Square	β	β	α

The next rule is termed a *conjunction/disjunction*. By this rule, one stimulus set is specified by a conjunction of one value on each of the relevant attributes (alpha is red and triangular). The other set is specified by a disjunction of the remaining values on each attribute (beta is green or square). Each relevant attribute in this concept is correlated with the responses, but the correlation is less than perfect. For example, if the subject attends only to color, he should get 75% of the stimuli correct because he can get all of the green stimuli correct but only half of the red stimuli.

The third rule is termed a *biconditional*. By this rule, each stimulus set is specified by a biconditional relationship between two attributes (alpha is red and triangular or not-red and not-triangular). Neither relevant attribute in this concept is correlated with the responses. For example, if the subject attends only to color he should get only 50% correct (chance) because half of the red stimuli and half of the green stimuli belong in each category.

The magnitude of these correlations depends on the particular distribution of the stimuli. For example, by the biconditional concept and with more triangles than squares, color would be correlated with the responses because most of the alphas would be red (triangles) and most of the betas would be green (triangles). However, the above examples suffice to demonstrate that the correlations between individual relevant attributes and the responses are dependent upon the conceptual rule and may vary from 0 to 1.

There are alternative ways of describing these three concepts. For example, the affirmation can also be described as a negation (alpha is not-green). In the conjunction/disjunction, one stimulus set can be described as an exclusion (alpha is red and not-square) and the other set can be described as an implication (beta is not-red or square). The biconditional can also be described as an exclusive disjunction (alpha is red or square but not both). Additional descriptions are possible (see Neisser and Weene, 1962), but with binary attributes all of these descriptions reduce to the three different response assignments shown in Table 1.

METHOD

—Concepts—The three concepts described above were used; affirmation, conjunction/disjunction, and biconditional. The distribution of stimuli was determined by the distribution shown in Table 1. Therefore in the conjunction/disjunction concept, the sets of stimuli associated with each response differed in size, with three-fourths of the stimuli in the disjunctive category. In the other two concepts, the stimulus sets were of equal size.

—Correlated irrelevant attributes—Each concept was presented under three degrees of correlation: either all irrelevant attributes were orthogonal to the responses, or one irrelevant attribute was correlated with the responses, or two irrelevant attributes were correlated. In the affirmation and biconditional concepts, it was possible to classify 75% of the stimuli correctly by using only one correlated irrelevant attribute. Since it was possible to classify 75% of the stimuli correctly in the conjunction/disjunction concept by attending to a single *relevant* attribute, the correlated irrelevant attributes were arranged so that 87.5% of the stimuli could be classified correctly by using only one correlated irrelevant attribute. Thus, in the conjunction/disjunction and biconditional concepts, the correlated irrelevant attributes were more highly correlated with the responses than were the relevant attributes.

—Stimuli and attributes—The stimuli were schematic pictures of 'bugs' that varied on six binary attributes: body shape, leg number, spot color, antenna type, eye type, and foot type. The relative saliency of the attributes was determined by pilot work, and the four most salient attributes were divided into two pairs. One pair consisted of body shape and spot color, the other of leg number and antenna type. The relevant attributes were selected from one pair and the correlated irrelevant attributes from the other. These selections were completely counterbalanced. The remaining attributes were irrelevant and orthogonal to the responses.

A set of 16 stimuli was selected for each condition. Each set was presented to the subject in four random orders. The attributes were orthogonal to *each other* in all conditions except the conjunction/disjunction conditions, where the correlated (with the responses) irrelevant attributes were also correlated with each other and with the relevant attributes.

—Procedure and instructions—The subject was told that he was in a learning experiment and his task was to learn to discriminate between two classes of bugs, 'alphas' and 'betas.' He was told nothing about the nature of possible solutions.

After the instructions were read, the subject was shown two stimuli that differed on all six attributes and was asked to list the differences between the stimuli. If he omitted an attribute, it was pointed out to him.

The learning task was then started. The stimuli were presented one at a time. When a stimulus appeared, the subject responded verbally. The experimenter provided feedback by giving the correct response immediately after the subject's response or after the stimulus had been displayed for 4 sec. After the set of 16 stimuli had been presented, the subject was asked to write a rule that would discriminate between the alphas and the betas. Guessing was encouraged. This procedure was repeated until he reached a criterion of one perfect classification of the set of 16 stimuli or until the set had been presented 20 times.

The stimuli were presented in the form of slides and back-projected. Each slide was presented for 5 sec, with about 1 sec between slides. The experimenter and the subject were in separate rooms and communicated with an intercom.

—Subjects—The subjects were paid male volunteers from the psychology subject pool at Johns Hopkins University. Eight subjects were run in each of the nine conditions (three concepts by three degrees of correlation), for a total of 72 subjects.

RESULTS

Trials to criterion

Table 2 shows mean trials to criterion (excluding the 16 criterion trials) for each of the three concepts as a function of the presence or absence of correlated irrelevant attributes. (Number of correlated irrelevant attributes had no effect, and data are pooled across conditions involving one and two correlated irrelevant attributes.) The three concepts differed in difficulty [$F(2, 66) = 42.8, p < .001$], and the order of difficulty was the same as that found by previous investigators (Neisser and Weene, 1962). Each concept was significantly different from the other two concepts [for each pair, $F(1, 66) > 4.25, p < .05$]. There was no effect due to the presence or absence of correlated irrelevant attributes ($F < 1$), and no interaction ($F < 1$).

Table 2. Mean trials to criterion

Type of irrelevant attributes	Concepts		
	Affirmation	Conjunction/disjunction	Biconditional
All orthogonal	8.6	171.4	198.9
One or two correlated	36.9	167.2	217.9

Within the affirmation concept, the variance was much greater when correlated irrelevant attributes were present [$F(15, 7) = 159.3, p < .001$]. This increased variance is due primarily to the 5, of 16, subjects who had scores more than five standard deviations above the mean of the condition without correlated irrelevant attributes. There was no change in variance in the other two concepts.

Error rate and error distribution

Table 3 shows mean error rate (total errors divided by trials to criterion, averaged across subjects) for each concept as a function of the presence or absence of correlated irrelevant attributes. The error rate differed across concepts [$F(2, 66) = 15.2, p < .001$]. The subjects in the groups with a conjunction/disjunction concept had a lower error rate than the subjects with the other two concepts [each $F(1, 66) > 17.5$, $p < .01$]. The error rates of the groups with the affirmation and biconditional concepts were not different [$F(1, 66) = 1.0$] from each other. The error rate was also lower when correlated irrelevant attributes were present than when they were absent [$F(1, 66) = 7.6, p < .01$]. The interaction was not significant ($F = 1.1$).

Table 3. Mean error rate

Type of irrelevant attributes	Concepts		
	Affirmation	Conjunction/disjunction	Biconditional
All orthogonal	.46	.35	.47
One or two correlated	.43	.22	.38

It seemed possible that the reduced error rate in the conjunction/disjunction concept might be due to a response bias, since the stimulus categories were not equally likely. On the other hand, if there were no response bias, one would expect 75% of the errors to occur on stimuli that belonged in the disjunctive category, since 75% of the stimuli were in this category. The actual proportion of errors on these stimuli was 74%, so it would appear that the reduced error rate was due to the correlation between the *relevant* attributes and the responses, and not to any response bias.

The reduced error rates when correlated *irrelevant* attributes were present suggest that the subjects attended to these attributes and based their responses, at least to some extent, on the values of these attributes. If this were true, errors would not be randomly distributed across stimuli. For example, if the subject attended to a correlated irrelevant attribute that allowed him to get 75% of the stimuli correct, more than one-fourth of his errors should occur on the remaining 25% of the stimuli. This distribution of errors was in fact shown by 14 of the 16 subjects in the affirmation concept (binomial test, $p < .01$) and by 16 of the 16 subjects in the other two concepts when correlated irrelevant attributes were present.

Precriterion hypotheses

The subjects' precriterion hypotheses are of some interest because if subjects attend to the correlated irrelevant attributes, these attributes should be mentioned more than the uncorrelated irrelevant ones. The specific hypotheses are not very useful, since guessing was encouraged, and they are often unintelligible. Furthermore, most subjects in the affirmation concept reached criterion so quickly that there are not enough precriterion hypotheses for an analysis. However, it is possible simply to tabulate the attributes mentioned by the subjects in the conjunction/disjunction and biconditional groups.

Since the four most salient attributes were completely counterbalanced over the roles they could play in any condition (relevant, correlated irrelevant, or orthogonal irrelevant), it is meaningful to examine the probability of an attribute being used in a precriterion hypothesis as a function of its assignment. The average probability of usage was .56 for a relevant attribute, .56 for a correlated irrelevant attribute, and .27 for an orthogonal irrelevant attribute. The average probability of usage of the orthogonal irrelevant attributes was significantly less than that of both the relevant attributes ($t = 6.35$, $df = 142$, $p < .001$) and the correlated irrelevant attributes ($t = 5.67$, $df = 94$, $p < .001$).

These precriterion hypotheses were not sequentially independent. An examination of the conditional probabilities that an attribute would be mentioned, given that it was mentioned in the preceding hypothesis, shows a significant increase over the overall probabilities for each of the different classes of attributes (for relevant attributes, $t = 2.96$, $df = 183$, $p < .01$; for correlated irrelevant attributes, $t = 3.16$, $df = 90$, $p < .01$; for orthogonal irrelevant attributes, $t = 3.46$, $df = 86$, $p < .01$). These conditional probabilities were .69 for relevant attributes, .75 for correlated irrelevant attributes, and .46 for orthogonal irrelevant attributes. The average conditional probability of usage of the orthogonal irrelevant attributes was significantly less than that of both the revelant attributes ($t = 3.72$, $df = 127$, $p < .001$) and the correlated irrelevant attributes ($t = 4.51$, $df = 82$, $p < .001$).

Final rules

In the conditions with the irrelevant attributes all orthogonal, the solutions involved rules using only the relevant attributes. In the other conditions, several additional solutions were possible, solutions involving the correlated irrelevant attributes. Every solution using correlated irrelevant attributes required the use of more attributes than did the solutions using only relevant attributes. Even so, an examination of the final rules written by the subjects who reached criterion shows that 16 subjects used correlated irrelevant attributes when they were available. Table 4 shows the distribution of these subjects across the conditions. More rules using correlated irrelevant attributes occurred in the groups with the conjunction/disjunction and biconditional concepts than in the group with the affirmation concept ($\chi^2 = 7.91, df = 1, p < .01$).

Table 4. Distribution of final rules using correlated irrelevant attributes

Number of correlated irrelevant attributes	Concepts		
	Affirmation	Conjunction/disjunction	Biconditional
One	0 (8)	4 (7)	1 (7)
Two	1 (8)	5 (8)	5 (6)

Note: Each cell entry is the number of subjects who formulated rules using correlated irrelevant attributes. The number of subjects to reach criterion is shown in parentheses. $N = 8$ subjects per condition.

The rules using correlated irrelevant attributes were of two basic types. In one type, the set of stimuli associated with a single response was divided into a number of subsets and each subset was described as a conjunction of two or three attribute values. In the other type of rule, the stimulus set associated with a single response was specified by a value of a correlated irrelevant attribute and then a few stimuli were listed as exceptions, with fairly specific descriptions involving conjunctions of at least three attribute values.

The rules that did not use correlated irrelevant attributes are also of some interest. Descriptions of the affirmation concept and of the conjunctive category of the conjunction/disjunction concept were always simply affirmations and conjunctions respectively. Nine subjects described the disjunctive category of the conjunction/disjunction concept. Only one of these subjects used a disjunctive rule. The remaining eight used a rule that divided the stimuli in the disjunctive category into two subsets. One of these subsets was specified by a single value on one of the relevant attributes, and the other subset was specified by a conjunction of the other value on the previously used attribute and a value on the other relevant attribute. For example (referring back to Table 1), the betas might be described as anything green plus the red squares.

There were also 12 subjects who reached criterion in the biconditional concept and did not use correlated irrelevant attributes. All of these subjects used a rule that divided the set of stimuli assigned to a single response into two subsets, each of which was specified as a conjunction of one value on each of the relevant attributes. For example (referring to Table 1), the alphas might be described as the red triangles plus the green squares.

DISCUSSION

Correlated irrelevant attributes as distractors

These results suggest that increasing the correlation between an irrelevant attribute and the responses increases the probability of a subject's attending to the attribute—and that this attention interferes with his discovery of the relevant attributes but produces a lower error rate. This interference has two different possible effects. It may simply increase trials to criterion, or it may lead to the development of a solution involving the correlated irrelevant attribute. In the groups with the affirmation concept, the correlated irrelevant attributes primarily produced an increase in trials to criterion. The effect was quite small and took the form of an increase in variance, since only about one-third of the subjects were led astray by the correlated irrelevant attributes. This small effect was probably due to the easily perceived direct relationship between the relevant attribute and the responses in this concept. In the groups with the conjunction/disjunction and biconditional concepts, the presence of the correlated irrelevant attributes did not change trials to criterion but did lead a large proportion of the subjects to develop a new rule, a rule primarily based on these attributes.

The results obtained by Gormezano and Grant (1958) and by Abraham, Gormezano, and Wiehe (1964) are consistent with this interpretation. They found that increasing the correlation between an irrelevant attribute and the responses increased trials to criterion but had no effect on total errors. They used a single-attribute concept slightly more difficult than the affirmation used in the present study, and this resulted in a significant increase in solution time. However, since attention to the correlated irrelevant attribute allowed a lower error rate, there was no change in total errors.

The results obtained by Bourne and Haygood (1960) are not consistent with the interpretation we propose. They found that increasing the correlation of an irrelevant attribute decreased both trials to criterion and total errors. However, they apparently gave their subjects detailed instructions on the nature of possible solutions, while neither the present study nor those by Gormezano and Grant (1958) and by Abraham, Gormezano, and Wiehe (1964) did. This difference in methodology could possibly account for the effects Bourne and Haygood obtained.

All-or-none versus incremental learning

The present results have some general implications for models of concept attainment. The mean error rate in the affirmation concept was .46 when all irrelevant attributes were orthogonal to the responses. This is consistent with the all-or-none models of attribute or hypothesis testing that have been developed for single-attribute concepts (see Bower and Trabasso, 1964). For the biconditional concept with all irrelevant attributes orthogonal, the mean error rate was .47; detailed analysis shows that it was stable near chance level and then sharply dropped when the subject reached criterion. This result is also consistent with an all-or-none model, and it suggests that the hypothesis-testing models could be extended to include biconditional concepts by some simple modification such as the assumption that the subject samples hypotheses involving two or more attributes after he has tried all hypotheses involving single attributes.

For the conjunction/disjunction concept and in all conditions involving correlated irrelevant attributes, however, it is apparent that an all-or-none model is inappropriate: error rates were below chance, errors were not randomly distributed across stimuli, and the precriterion hypotheses were not sequentially independent. In all of these conditions there were attributes (relevant or irrelevant) correlated, but less than perfectly so, with the responses. The data imply that when a subject finds a hypothesis allowing above-chance but less than perfect performance, he stops sampling new hypotheses and tries to perfect the partially correct one.

Taken together, these results suggest that subjects start by sampling and testing hypotheses in an apparently random manner. The selection of hypotheses is probably not random but based on some criterion—such as attribute salience—that is irrelevant to the concept and thus results in apparently random responding. If the subject finds a hypothesis that is partially correct, he keeps the hypothesis and tries to modify it. Thus, an all-or-none model seems appropriate for situations where no simple partially correct hypothesis exists. In situations where such a hypothesis does exist, an all-or-none model is inappropriate.

It should be noted that this analysis suggests that the subject's attention is focused on the attributes which are correlated with the responses. Thus, multiattribute concepts in which the relevant attributes are correlated with the responses will be easier than multiattribute concepts in which the relevant attributes are orthogonal, because the correlations will lead the subject to focus on the relevant attributes. In the present study, therefore, the difference in difficulty between the conjunction/disjunction and the biconditional concepts could have been due to the relative ease with which the relevant attributes could be identified in the conjunction/disjunction concept. If the subject's attention could be focused on the relevant attributes by some other means (such as by using very salient relevant attributes), the difference in difficulty of these two concepts might well disappear.

Final rules

As was mentioned in the introduction, each of the stimulus sets assigned to a response could have been described by several different rules using only the relevant attributes. For example, the conjunctive category of the conjunction/disjunction rule could have been described as a conjunction, exclusion, or conjunctive absence (using the terminology of Neisser and Weene, 1962). The disjunctive category could have been described as a disjunction, implication, or disjunctive absence. Either category of the biconditional concept could have been described as a both/neither (biconditional) or an either/or (exclusive disjunction). In the present study, except for the affirmation and conjunction rules, almost none of these logical possibilities was used. Furthermore, the rules that were used were almost all of the same form: except for the single disjunctive rule, *all* of the rules (including those using correlated irrelevant attributes) described the set of stimuli assigned to a response by listing one or more attribute values required in the stimuli. If the entire set of stimuli could not be described in this manner, the stimuli were divided into subsets that could be so described. And, of course, such a description that uses only one attribute is formally an affirmation, and one that uses more than one attribute is formally a conjunction.

It is well known (Bruner, Goodnow, and Austin, 1956; Conant and Trabasso, 1964) that people have some difficulty attaining disjunctive concepts. In most studies the conceptual rule is specified beforehand; in the present study the conceptual rule was not specified, and the disjunctions and other complex concepts were actually learned as combinations of affirmations and conjunctions. It is not clear how far this finding may be extended. For example, if the number of values per attribute were increased, the proportion of disjunctive rules used by the subjects would probably increase. However, it is worth noting that the rule the subject uses to solve the concept is not necessarily the one the experimenter had in mind. The kinds of rules that are learned under various experimental conditions would be a fruitful topic for further investigation.

Notes

This article is based on a dissertation submitted to the faculty of the Johns Hopkins University in partial fulfillment of the requirements for the doctoral degree. The author wishes to acknowledge the helpful comments and criticisms offered by Howard Egeth and W. R. Garner. Received for publication September 24, 1970.

References

Abraham, F., Gormezano, I., and Wiehe, R. 1964. Discrimination learning as a function of prior relevance of a partially reinforced dimension. *Journal of Experimental Psychology* 67:242–249.

Bourne, L. E., Jr., and Haygood, R. C. 1960. Effects of intermittent reinforcement of an irrelevant dimension and task complexity upon concept identification. *Journal of Experimental Psychology* 60:371–375.

Bower, G. H., and Trabasso, T. 1964. Concept identification. In *Studies in mathematical psychology*, ed. R. C. Atkinson, pp. 32–94. Stanford, Calif.: Stanford University Press.

Bruner, J. S., Goodnow, J. J., and Austin, G. A. 1956. *A study of thinking*. New York: Wiley.

Conant, M. B., and Trabasso, T. 1964. Conjunctive and disjunctive concept formation under equal information conditions. *Journal of Experimental Psychology* 67:250–255.

Gormezano, I., and Grant, D. A. 1958. Progressive ambiguity in the attainment of concepts on the Wisconsin Card Sorting Test. *Journal of Experimental Psychology* 55:621–627.

Neisser, U., and Weene, P. 1962. Hierarchies in concept attainment. *Journal of Experimental Psychology* 64:640–645.

American Journal of Psychology
1971, Vol. 84, No. 3

THE APPARENT LENGTHS OF PERIMETERS AND DIAMETERS DEFINE A RATIO SMALLER THAN PI

Robert Teghtsoonian and Martha Teghtsoonian
Smith College

Fourteen observers made magnitude estimations of the apparent length of the diameters and perimeters of contour circles. The ratio of apparent perimeter to apparent diameter was only about three-fourths of the value of π, the ratio of the physical counterparts.

There are a number of reports on scales of apparent length showing apparent length to be a power function of physical length with an exponent near 1.0.[1] In these studies the observers were asked to judge rectilinear extent, and we have found no evidence that the resulting scales apply to curvilinear extent as well. To explore this problem we presented contour circles, varying from 2.8 cm to 30.0 cm in diameter, to a group of 14 observers and obtained their magnitude estimations of the apparent length of the diameters and perimeters. For each observer, half the stimuli presented were accompanied by a request for a judgment of the diameter and the other half by a request for a perimeter judgment. The two judgments were intermixed according to a random sequence to ensure that each observer employed the same scale in making both varieties of judgments. No standard length was provided, nor any modulus suggested; thus the resulting multiplicative constants, as well as the exponents, could be meaningfully compared.

For each attribute a straight line was fitted by the method of least squares to the log geometric means of the judgments as a function of log length in inches. The resulting power functions are:

$$\text{for diameters, } \psi_D = 1.54\, D^{1.09}\,;$$
$$\text{for perimeters, } \psi_P = 1.13\, P^{1.07}\,.$$

The exponents are nearly identical and are typical of values reported previously for similar ranges of linear extents. The interesting difference between the two functions is in the multiplicative constants, or scale factors: a diameter of given length appeared longer than a perimeter of that same length. If the simplifying assumption is made that both exponents are equal to unity, then the following relations hold:

$$\psi_P = 1.13\,P^{1.0}\,,$$
$$\psi_D = 1.54\,D^{1.0}\,, \text{ and}$$
$$\psi_P/\psi_D = (1.13/1.54)(P/D) = .735\,\pi\,.$$

This is to say that the ratio of apparent perimeter to apparent diameter is only about three-fourths the value of π, the ratio of the physical counterparts.

If, as we have shown, perimeters are underestimated relative to diameters, it may be true that, in general, curvilinear extents are underestimated relative to rectilinear equivalents. For example, the perimeter of a cylinder is commonly underestimated relative to its height. A simple model can easily be constructed to illustrate this illusion: join two parallel edges of a square piece of paper to form a cylinder. Most adults will judge the height to be greater than the perimeter, although they are in fact equal.

Notes

The authors thank Mary Viola for her help in preparing stimulus materials, running subjects, and tabulating data. Received for publication March 3, 1970. Edited by Professor E. B. Newman.

1. For a summary of these reports and an additional example, see M. Teghtsoonian and R. Teghtsoonian, Seen and felt length, *Psychonomic Science*, 1965, 3:465–466.

NOTES AND DISCUSSION

THE SIXTY-SEVENTH ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The sixty-seventh annual meeting of the Society of Experimental Psychologists was held at the University of Colorado, Boulder, on April 9 and 10, 1971. Gregory A. Kimble, Chairman of the Society for the year, presided at the business meeting and at the sessions in which scientific papers were presented.

The following Members and Fellows were present: DeValois, Duncan, Grant, Grice, Hirsh, D. Hurvich, L. Hurvich, Hunt, Kimble, Krech, Liberman, Logan, Mackworth, Meyer, Ratliff, Solomon, Stellar, Teitelbaum, Thomas, Volkmann, Wickens.

Fred Attneave, Robert Boynton, Richard Held, Michael Posner, Burton Rosner, and Saul Sternberg were elected to membership in the Society.

Fellow John Paul Nafe and Charter Member Harry P. Weld died during the year. The present membership in the Society is 78 Members and 40 Fellows.

The 1971 Warren Medals were awarded to Leo M. Hurvich and Dorothea Jameson Hurvich, "For studies that have given quantitative substance to the Hering model of opponent processes in the visual system."

The Society accepted the invitation of Stanford University to meet in Stanford, California, in 1972. Ernest R. Hilgard was elected Chairman for 1971–72. The Executive Committee consists of the Chairman, the Secretary, Ira Hirsh, and Howard Kendler. The Trustees are Frank Geldard, Daniel Lehrman, and the Secretary.

Carl P. Duncan, Secretary; *Northwestern University*

THE SIXTY-THIRD ANNUAL MEETING OF THE SOUTHERN SOCIETY FOR PHILOSOPHY AND PSYCHOLOGY

The sixty-third annual meeting of the Southern Society for Philosophy and Psychology was held on April 8–10, 1971, at the Center for Continuing Education at the University of Georgia, Athens, Georgia. Approximately 250 members, friends, and guests were in attendance. The program consisted of three symposia in psychology, one symposium in philosophy, and two joint sessions. Altogether there were 35 papers in philosophy and 56 papers in psychology. There was also a festscrift program in honor of Professor Frank Geldard. Tours of the University of Georgia Psychology Building were provided for the membership. As usual, a social hour and business meeting were also included in the program.

At the first of the two joint sessions, Malcolm Arnoult, Texas Christian University, introduced Dwight Van de Vate, University of Tennessee, who delivered the presidential address, "Notes toward a Theory of Time." At the second joint session, William Bevan, Executive Officer of the American Association for the Advancement of Science, delivered an invited address, "Higher Education in the Seventies—A Once and Future Thing." Discussants were Earl A. Alluisi, University of Louisville (psychology), and John Post, Vanderbilt University (philosophy).

Two Junior Awards (now the Richard M. Griffith Memorial Awards) were given. The recipient in philosophy was Edward S. Shirley, Louisiana State University, for his paper "Stimulus Meaning and Interdeterminancy of Translation." The recipient in psychology was George W. Menzer, University of Louisville, for his paper "Further Experiments with Orientation-specific Chromatic Aftereffects." These awards are given annually for the papers of highest merit presented by members who have not yet received the doctoral degree or who have received it within the last five years.

Cochairmen of the Committee on Local Arrangements were John T. Granrose (philosophy), University of Georgia; and Clyde E. Noble (psychology), University of Georgia. Cochairmen of the Program Committee were Nancy D. Simco (philosophy), Memphis State University; and Joel S. Warm (psychology), University of Cincinnati. Cochairmen of the Junior (Griffith Memorial) Awards Committee were Maynard Adams (philosophy), University of North Carolina; and Malcolm

Arnoult (psychology), Texas Christian University. At the annual business meeting, reports of the officers and of the several standing and ad hoc committees were received and accepted. It was announced that the sixty-fifth annual meeting would be held in Knoxville, Tennessee.

The deaths of members John R. Binford, Elizabeth Duffy, Theodore Lafferty and Charles Odum were noted with sorrow. Brief memorials were read for Elizabeth Duffy and Theodore Lafferty, and for James Simmons, decreased in 1969.

There were 31 philosophers and 36 psychologists elected to full or associate membership. Michel Loeb, University of Louisville (psychology), and Douglas Browning, University of Texas (philosophy), served on the membership committee. Elected as *Members in Philosophy* were Linda A. Bell, Marcus Clayton, William F. Edwards, William H. Friedman, Alan Gibbons, R. Kirby Godsey, Iredell Jenkins, William J. Kilgore, Douglas M. McDonald, Veronica A. Miller, Neale H Mucklow, Thomas E. Pauly, Edward S. Shirley, Dieter Turck, and Robert H. Vorsteg. Elected as *Members in Psychology* were Lloyd Hastings, Robert H. Pollack, Katherine Vickery, Robert M. Stutz, C. Ray Carpenter, Jim Larimer, R. J. Senter, Ernest M. Weiler, John D. Seago, Thomas Mulhern, Henry E. Marks, Paul Cameron, Janet F. Rees, Kevin M. Mitchell, Anthony F. Grasha, George B. Greaves, Earl C. Brown, Frank C. J. McGurk, Janice R. Teal, Avrum I. Silver, Marsie M. Field, and Harold L. Hawkins. Elected as *Associate Members in Philosophy* were Gary J. Acquaviva, James B. Davant, Frank K. Fair, Terry M. Goode, Jerry Gravander, David R. Hiley, William E. Hoffmann, Bart F. Kennedy, Charles W. Lewis, Jr., James J. Meyer, Robert E. Majzler, S. J., Frederick Michael, Gale A. Townsley, Stephen W. White, James E. Abbott, and James C. Edwards. Elected as *Associate Members in Psychology* were David J. Micco, Jr., Jeffrey L. Clark, Michael F. Sherrick, Robert L. Brunner, Ronald R. Rossi, Julius J. Persensky, Florrie M. Pennington, Johnny K. Worley, Howard L. Bregman, Claude R. Miller, Jr., G. Rufus Sessions, Ernest H. Harrell, Stephen H. Hobbs, and Franklin M. Berry.

The following officers were elected to serve during 1971–72: Stanley B. Williams, College of William and Mary, President; Douglas Browning, University of Texas, President-Elect, William T. Blackstone, Secretary, University of Georgia; Carr Payne, Georgia Institute of Technology, Treasurer; Dwight Van de Vate, University of Tennessee, Past President. Continuing members of the Council include Donald S. Lee, Tulane University; Lelon J. Peacock, University of Georgia; Herbert Kimmel, University of South Florida; and Ramon Lemos, University of

Miami. Clyde Noble (psychology), University of Georgia, and Donald Sherbourne (philosophy), Vanderbilt University, were elected for three-year terms.

The sixty-fourth annual meeting of the Southern Society for Philosophy and Psychology will be held at the Sheraton-Jefferson Hotel, Saint Louis, Missouri, on March 30–April 1, 1972. It will be hosted by the Department of Psychology of Washington University.

W. T. Blackstone, Secretary; *University of Georgia*

THE 1971 ARROWHEAD CONFERENCE

For the eighth year, a small group of investigators concerned with problems in verbal learning (broadly conceived) met for an informal exchange of ideas on June 14–18, 1971. This year, as in most previous years, the group convened at the University of California Conference Center, Lake Arrowhead, California. In attendance were R. A. Bjork, G. Bower, R. G. Crowder, D. L. Hintzman, E. Martin, A. W. Melton, B. B. Murdock, Jr., R. W. Schulz, E. Tulving, M. T. Turvey, B. J. Underwood, and D. D. Wickens. The conference is supported by NIMH Grant MH 18418 to the Ohio State Research Foundation, D. D Wickens principal investigator.

One session, a general one, dealt with what some believe to be serious problems of terminology in the field. No agreeable plan for resolving these problems was found. In the remaining sessions, each investigator presented his current experimental work, or his current theoretical thinking, or both. What follows is a brief summary of these remarks. Further details (if not the remarks themselves in any 'final' form) may be obtained by writing directly to the participants.

Bjork's report was concerned in general with the effects of serial position on performance in immediate and delayed tests of free recall and recognition, and in particular with the role played by subjects' control-process activities in producing those effects. He considered the possible causes of the 'negative recency' effect in delayed retests of free recall in some detail, especially the possibility that the spacing between the input and output of an item during immediate recall may determine the extent to which the output event facilitates later attempts to recall the item. Bjork argued that serial-position effects result from what subjects choose to do, and are not unavoidable consequences of associative interference, temporal saliency, or other intrinsic factors.

Bower marshalled the arguments against simple one-process theories that recall and recognition are mediated by different thresholds on the strength of a single memory trace. Similarly, he marshalled detailed arguments against the simple strength or familiarity theory of recognition memory, a theory which is contradicted by several elementary facts about recognition and whose extensions to recency judgments and frequency-of-experience judgments appear misguided. In place of that theory, Bower proposed that recognition is mediated by retrieval of information about contextual events that surrounded the presentation of the item in question and were associated with the item when it occurred. Assuming gradual drift over time in the contextual milieu and admitting the possibility that each repetition of an item establishes an independent association to a new pattern of contextual elements, the proposal provides the mechanisms for a theory about recognition memory, list differential, list-editing of recall, recency judgments from memory, frequency-of-experience judgments from memory, and lag effects on frequency judgments.

Crowder reported three sets of data on the properties of a precategorical acoustic-storage system. The first group of studies showed that such a memory store is completely insensitive to semantic organization but is affected by physical similarity. The second group of studies explored the temporal-decay property of precategorical acoustic storage. Finally, recent data indicated conspicuously different acoustic memory results for vowels and consonants.

Hintzman reported on several experiments using memory judgments to determine what information is stored about experimental presentations of a word. Two experiments attempted to show that a repeated word has more than one 'time tag,' and that this may be the way frequency is represented in memory. A third experiment showed that subjects retain input modality of a word for at least several minutes after presentation. A fourth showed that if a word is repeated in different modalities, subjects can later tell which modality came first, suggesting that the 'time tag' and 'modality tag' for a particular presentation are tied together in memory.

Martin presented data demonstrating that components of a compound stimulus act independently and do not form intercomponent associations during paired-associate learning. Also reported were preliminary serial-learning results indicating that *backward* (as opposed to *forward*) assembly of parts into a whole produces faster combination-phase learning and better retention over 24 hours.

Melton reported studies providing further evidence on the relations between procedural variables and the 'lag effect' in free recall. Particular attention was paid to variables that may be responsible for contradictions in the findings from different laboratories. In addition, he discussed the results of some exploratory experiments in terms of the multiple-encoding interpretation of the lag effect.

Murdock's remarks centered on recent modality studies that suggest separate prelinguistic auditory and visual stores with persistences of at least several seconds. These stores show differential interference effects. He also compared a strength theory and a finite-state model in accounting for data from (*a*) a recognition-memory study with confidence-judgment and latency measures, and (*b*) a judgment-of-recency study showing a crossover effect between CCCs and CVCs as a function of lag.

Schulz focused his presentation on work concerned with (*a*) the development of proactive inhibition in free recall in relation to Underwood's theory on the attributes of memory; (*b*) transfer under recognition as opposed to recall conditions; (*c*) transfer in the A-B, A-B′ paradigm; (*d*) a subject's ability to discriminate in the mediation paradigm between facilitation and interference pairs under mixed-list conditions of presentation.

Tulving's presentation was an attempt to persuade other participants of the conference that a heuristically useful distinction may be drawn between (*a*) episodic memory, memory for temporally dated personal experiences, and (*b*) semantic memory, memory for meanings of concepts, relations, and rules for manipulating symbols and concepts. He also pointed out that almost all research in human memory and verbal learning in the past has been concerned with phenomena of episodic memory.

Turvey pointed out that in vision, backward masking by pattern originates either peripherally or centrally. When the perceptual impairment is of peripheral origin, a simple relation exists between the target stimulus and the minimal time needed to evade the masking action: target-stimulus energy × minimal interstimulus interval = a constant. When the masking is of central origin, energy variables are irrelevant; what matters is the time elapsing between onset of the target and onset of the mask. Patterns that mask peripherally may not necessarily mask centrally; moreover, forward masking is primarily of peripheral origin. He also discussed both data and theory in the context of techniques to investigate processing of linguistic information in vision.

Underwood discussed preliminary results from experiments dealing with the massing and distribution of multiply occurring words in free recall. The superiority of distributed items was found with varying rates of presentation, with concrete and abstract words, and with imagery instructions. Recall of items occurring four times within a list was found to be directly related to the number of spacings. In a further experiment, subjects paced their study time for each presentation of a word. Massed items received less direct study time than did distributed items, but the difference was insufficient to account for the recall differences.

Wickens covered four topics. First, he reviewed current knowledge of variables mediating 'PI release' in short-term memory. Second, he reported a retroactive-inhibition study in which a large positive effect on recall was demonstrated when a consistent contextual activity preceded both original learning and recall. Third, he discussed evidence suggesting that for some characteristics of words the encoding occurs prior to the perception of the word. Finally, he pointed out that the free recall of response terms following A-B, A-D learning shows a form of 'spontaneous recovery' as a function of the length of the interval following A-D learning.

Benton J. Underwood; *Northwestern University*

BOOK REVIEWS

Julian Hochberg
Columbia University

Perceptrons

By Marvin Minsky and Seymour Papert. Cambridge, Mass.: The M.I.T. Press, 1969. Pp. 258. $12.00; paperback, $4.95.

Psychologists have long been interested in the idea of self-organizing nets of neurons, an interest sparked both by the McCulloch-Pitts model of logical elements in neurons and by Hebb's model of learning. A major step forward in the study of such networks occurred when Frank Rosenblatt proposed the *Perceptron* (The Perceptron, a probabilistic model for information storage in the brain, *Psychological Review*, 1958, 65:386–418; and *Principles of neurodynamics*, Washington, D.C.: Spartan Press, 1962).

Roughly, the Perceptron is a Hebbian-like model of a self-organizing system. It was actually created both as a computer program and as a physical machine, and a number of experiments in discrimination learning were conducted with it. Basically, it contains three sets of neuronlike elements. The first set is the *retina*, a set of on-off elements. As the name implies, this set is analogically similar to the retina of the eye. It is best thought of as a receptor field. Retina elements are connected to *association* elements, either randomly or in some design. The association elements, instead of being on-off elements, produce weighted input to a summing and threshold-detecting *response* unit. If the total input exceeds a threshold, the response unit fires.

The Perceptron is self-modified by a *training rule*, a function which changes the weights of the association element depending upon the output of the response unit. Thus if the response unit gives the 'proper' response to a picture projected on the retina, the association weights of active associators may be increased; or they can be decreased in response to improper outputs. A number of other training rules can be imagined. There are a number of variations on this theme. In a *cross-coupled* Perceptron the association elements' input is to each other as well as to the response element, while in a *back-coupled* Perceptron the association elements also transmit back into the retina unit.

Studies of Perceptron learning produced impressive results. The device was capable of learning a discrimination between letters projected on the retina. Cross-coupled and back-coupled Perceptrons could learn sequential responses at a primitive level. And although most of these points were made by demonstration, argument, and speculation, some mathematical proofs were also offered. A particularly interesting one was the proof that a simple Perceptron with a simple training rule can be constructed to learn any discrimination within its response repertoire, no matter what its initial responses are. This makes a Perceptron look very much like a neural model of learning, so the experimental studies were coupled with speculation about whether or not learning machines were just around the corner. Subsequent experience showed that they were not.

This reviews the status of things as of about 1965. At this time Minsky and Papert began their work. They realized that there were two unanswered questions about the Perceptron. What, mathematically, is the basic idea behind it? And if a Perceptron can learn anything that it can perform, what can it perform? These questions are answered in their book.

Their response to the first question has been to define a simple Perceptron that captures the basic idea behind the more complex models. To them a retina is a field of on-off elements. One can think of this as a vector, but it is probably better to think of it as a field of binary elements laid out on a plane—as a digitized picture. A picture on the retina is a designation of which units are on or off. (Throughout this review the terms 'on,' 'off,' and 'true,' 'false,' and '0,' '1' are used interchangeably.) A predicate, α, is a logical function on this field. In the original Perceptron terminology, a predicate corresponds to the connections that are needed to turn an association element on. To psychologists a predicate is simply a feature detector; to a mathematician it is a logical function of a sub-field of the retina. For instance, you could produce a predicate that detected a horizontal line in a particular position on the retina. An association is a weight, a_i, associated with predicate α_i. If α_i is true, then the quantity a_i is added to the input to the response element; otherwise it is not. Note that a_i could be negative. A Perceptron is a device consisting of m predicates and their associated weights. More formally, a Perceptron is a device that is capable of learning any discrimination that can be expressed by the function

$$\sum_{i}^{m} a_i\alpha_i \geq \theta$$

for a defined set $\{\alpha_i\}$ of predicates.

Part of the second question about the Perceptron is how the definition of the predicates determines the device's performance capability. Let the *order* of a Perceptron be the number of retinal elements connected to any individual predicate. For example, an order-one Perceptron has a weight assigned to each retinal element, but no weight assigned to any combination of elements, while an order-two Perceptron has weights assigned to all single elements and all combinations of two elements, and so on. Minsky and Papert show that an order-one Perceptron is capable of constructing discriminations logically equivalent to any of the 16 binary predicates on two arguments (retinal elements) *except* exclusive disjunction or equivalents. To those familiar with the concept-learning literature, this will be interesting, because it is well known that the exclusive-disjunction and equivalence discriminations are much harder for people to learn than the other binary discriminations. They also show that the concept of parity (which in their case means the question of whether or not there is an even or odd number of elements in an 'on state' on an arbitrarily large retinal field) cannot be learned by a Perceptron of order one, although this discrimination can be made by an almost trivial Turing machine.

Another class of performance results relates to transformation invariance. If we consider the retina as being a digitized picture on a two-dimensional field, what sort of transformations on this field can be tolerated if the device is to discriminate between two classes of objects? For example, from human experience we know that the visual field is almost impervious to translation and rotation. Minsky and Papert define a characteristic of *classes* of pictures which they call the *spectrum*. The spectrum of a class of pictures determines the order of the Perceptron needed to retain class invariance over translation and rotation. They next consider a more complex class of distortions, the topological transformations. These stretch the metrics of a space but retain the 'inside-out relations' between sub-pictures. If topological invariance is required, Perceptrons can only discriminate between classes of pictures that have different *Euler numbers*, where the Euler number of a picture is defined as the number of picture components less the number of holes in components. This result may be of less psychological interest, because it is not at all clear to me that humans can discriminate classes over topological transformations.

Finally, Minsky and Papert offer a number of new insights into the Perceptron learning theorems. Their major offering is that the Perceptron will stabilize its discrimination technique if perfect discrimination is not possible.

The psychologist who reads this book is going to have to work. The argument is mathematical and compares (favorably or unfavorably depending upon whether you are a mathematician or psychologist) with the more closely reasoned arguments in the *Journal of Mathematical Psychology*. Perceptron enthusiasts may complain that Minsky and Papert have considered only those Perceptrons that are mathematically tractable.

For the reasons just mentioned, one must be imaginative in evaluating Minsky and Papert's studies. Is it possible that the discriminations that can be learned by a Perceptron can be learned by any being that does not speak? In other words, is it language that makes us capable of learning concepts such as parity? I think that an interesting review of animal and human learning could be written relating the capabilities of different species to learn different discriminations with the order of Perceptron required to learn them. We might be able to develop some relation between the complexity of an intelligence-test item and the restrictions that would have to be placed on the Perceptron that could answer it. Such things might be done, but they have not yet. This may be an extremely important book that will trigger a major contribution to psychological theory. On the other hand, it may be a precise mathematical note on a rather obscure problem. Only time will tell.

Earl Hunt, *University of Washington*

The Nature of Psychology: A Selection of Papers, Essays, and Other Writings

By the late Kenneth J. W. Craik. Edited by Stephen L. Sherwood. New York: Cambridge University Press, 1966. Pp. xx, 184.

Three years before his death, Craik had published *The Nature of Explanation* (Cambridge: The University Press, 1943), a book well known to psychologists. It introduced the picture of man as a device that models events in the environment, and it argued for the deep relevance to psychology of such mathematics as the arithmetic of numerical series representing events in time and space. In that first book, Craik wrote: "Confusion of two similar concepts is a fruitful cause of erroneous thinking. I am particularly addicted to it, and if ever I conceive any original idea, it will be because I have been abnormally prone to confuse ideas, but have just saved myself by experimental verification, and have thus found remote analogies and relations which others have not considered!" (p. 92). The present book continues what his earlier book began, sketching out a design he did not live to complete.

K. J. W. Craik was a giant on whose shoulders we should dare to try to stand. The present book is a fiery posthumous collection of unfinished, unpolished, and thus undiluted notes, a feast of experimental and theoretical innovation. Craik died in his thirty-second year, five years past his Ph.D., a daring, inventive, and sometimes muddleheaded psychologist of towering genius. Parts of his work echo through the quarter century since his death to the psychological literature of our time.

The book begins with a reprinting of F. C. Bartlett's obituary notice, in which he evaluates Craik as an ex-student and sometimes colleague (*British Journal of Psychology*, 1946, 36:109); this is followed by Craik's epilogue to his 1940 unpublished doctoral dissertation, a discussion of the "marked difference of approach to perceptual problems in psychology and philosophy on the one hand and in physiology and physics on the other." In that statement he clears the decks for action, having decided that one should ask "how far the unknown (sensory qualities) can be explained in terms of the known (physiological structure and physical mechanisms)" and freed himself to suppose that "the reality of the physical world is not to be rashly denied" (p. 6).

All the rest of the book is fragments of uncompleted work. The first section (pp. 7–90) contains seminal statements relating learning and perception to servomechanisms and cyclic events in the nervous system. This work, of course, has been carried forward (see Iberall and McCulloch, *Currents in modern biology*, 1968). Particularly beautiful analyses by Craik include his analyses of neural networks and his serious attempts to introduce differential and integral calculus into the theoretical psychology of perception and thinking. For the former, he bases neural-net design on the method of Barrow's triangles, showing the analogy between differentiating a curve and the taking of successive differences, reducing the problem to the design of systems for performing successive subtraction (pp. 41–44). For the latter, he argues that "perception differentiates; thought integrates and extrapolates" (p. 63). Throughout this section he addresses himself to important thinkers of his time, particularly Hull, Thorndike, and Lashley. But he was after even bigger game, the foundation of a mathematical psychology that could use the power of the calculus. It is interesting that he neglected the work of Ernst Mach, who also tried.

He ignores Mach even in his notes on the discovery of an illusion in the perception of brightness (pp. 94–97). The illusion has been rediscovered more recently by Cornsweet (in Ratliff, *Mach bands*, 1967) and by Edwin Land (*Journal of the Optical Society of America*, 1971), but Craik's is one of the earliest discussions of the importance for brightness perception of the form of visual boundaries.

Other experimental work includes the startling report of Craik's visual observations of the temporal course of a foveal scotoma he induced by staring at the noonday sun for two minutes (pp. 98–103). Isaac Newton and Gustav Fechner had also dared to burn their eyes by staring at the sun, but Craik was slightly more circumspect, limiting his scotoma to one eye and analyzing the development of its recovery over months of time, noting the aniseikonia associated with it and the effects on it of flickering light. This study led him to believe that, in general, intermittent stimulation "is probably the best means of attaining high subjective brightness, preventing suppression and encouraging fixation," and to think that intermittent stimulation might prove useful in the treatment of amblyopia (pp. 104–105).

In this book we also see Craik in a more philosophical guise. In addition to theoretical and empirical notes are tidbits like this one: "We may be unable to make ourselves immortal, but if we write a drama about the thoughts and feelings of immortal gods, we shall have created something, something which may give us greater insight into our own place in the world, and the place of that world among abstract possibilities, and which at the same time transcends that world" (p. 178).

The selections in this book were made by its editor and by Warren S. McCulloch and Leo Verbeek. They are to be commended for including selections that convey a great deal of the mind of one of psychology's great men.

Herbert F. Crovitz, *Veterans Administration Hospital, Durham, North Carolina*

Selective History of Theories of Visual Perception, 1650–1950

By Nicholas Pastore. New York: Oxford University Press, 1971. Pp. 454. $10.00.

In reading this book one is vaguely reminded of an apochryphal tale of an art professor whose didactic technique was to take one painting from his favorite artist each semester and analyze it in minute detail. A student once asked him about this method of teaching and he replied: "It's true that I don't teach everything, but the part that I deal with I teach well."

The word 'selective' appears in the title of the book and warns the reader that all of the territory subsumed under the rubric of "Theories of Visual Perception" will not be covered. The fact is that it is very selective indeed, basically covering the empiricist viewpoint, with only the Gestaltists and Condillac as the opposition. Still, Pastore is scholarly and thorough in his treatment of the theorists he does deal with. His writing is generally clear, concise, and readable, with a few exceptions.

The book starts rather abruptly with an introduction that gives only a few clues as to what is to come and that fails to specify the rationale behind the choice of theorists or theoretical orientation. The opening chapters concern Descartes and Malebranche, neither of whom manifests himself as particularly important in terms of shaping the theories that follow. The next chapter opens the discussion of the empiricists with a treatment of Locke and Molyneux, followed by a fine presentation and analysis of Berkeley. Condillac, whose work is, unfortunately, not very well known to contemporary English-speaking psychologists, raised several objections to Berkeley's position. His criticisms serve as one of the few nativist islands in the text. The next two chapters are a montage of six theorists: Reid, Hamilton, Müller, Mill, Bain, and Lotze. They are presented briefly with more of an emphasis on similarities than on differences. Helmholtz's work on color contrast and space perception are discussed next. William James is presented with a great deal of care and reasonable detail.

The first section is then followed by a needed respite in the flow of theorizing. An attempt is made to review the theories and data assembled up to this point in time. The overview is helpful in placing the succeeding material in perspective. One of the best chapters in that work is the chapter on Gestalt psychology. The organization and treatment are commendable, despite the obvious absence of a discussion of the precursors to this theoretical stand.

The remaining two chapters are rather strange. In the first, Hull is termed a perceptual theorist and teamed with Hebb under the title of "connectionism." Although the treatment of Hebb is clear, the point of placing him with Hull never becomes apparent. The last chapter treats Ames and Murphy. The transactional viewpoint is primarily expounded via analysis of the Ames demonstrations. Murphy's inclusion is puzzling in light of the theories discussed above, and seems rather like an afterthought.

The book ends even more abruptly than it began, without any overview or summary analysis of the theoretical viewpoints presented. An interesting pair of appendixes is included, in which the Cheselden and the Wardrop cases of recovery from early blindness are presented in their original form. They are pleasant to read.

It should be clear from this overview that the book's major problem is its selectivity. It is true that the field of visual perception is probably too broad to be adequately surveyed in any one volume, but it is hard to rationalize any selection procedure that leaves out such major figures in perceptual theory as Hering, Wundt, Brentano, Stumpf, Fechner, Titchener, Munsterburg, Washburn, Helson, Werner and Wapner, Brunswick, and Gibson, to name but a few, and yet includes a chapter on Malebranche, about whom the author notes that "the invocation of divine intervention for explaining the mundane facts of seeing objects was essential. . . ."

This book, then, serves as a scholarly and useful survey of the empiricist viewpoint in perception, yet it is a shame that the author has limited his treatment to such a narrow selection of the many exciting viewpoints that have found expression between the years 1650 and 1950.

Stanley Coren, *New School for Social Research*

Fundamentals of Visual Science

By Melvin L. Rubin and Gordon L. Walls. Springfield, Ill.: Charles C Thomas, 1969. Pp. xxi, 435. $17.50.

As M. L. Rubin states, "much of the inspiration and material for the book which follows is from a lecture course in Physiological Optics presented at the University of California by my close, late friend and mentor, Gordon Lynn Walls" (p. xiii). Thus, although Rubin wrote the book, Walls is listed as co-author, and with good reason, since many of the book's virtues are due at least in part to Walls's vast knowledge of the field of vision and his concise, insightful analysis of its problems.

The book is divided into four parts, or units. The first, "Light and Basic Concepts," includes discussions of light, the measurement of light, and entoptic phenomena. The second, "Intensive Vision and Electrophysiology," concerns visual acuity and the photochemistry and electrophysiology of vision. The third, "Chromatic Vision," covers chromatic vision, with half its material devoted to color blindness. Finally, "Spatial Vision" has discussions of binocular 'fusion,' localization, and visual movement.

Although *Fundamentals of Visual Science* is primarily designed as a text to introduce the field of physiological optics to ophthalmologists, a book covering the above topics could easily find use as a text for an undergraduate or graduate course in sensory psychology or as a reference book for sensory psychologists engaged in visual research. However, I would definitely not recommend this book for either use. My main reason is simple: the book is badly out of date; only token attention is given to research done after 1960. As an example, perhaps extreme, a large portion of the material presented in the discussion of the electrophysiology of vision can be found in R. Granit's *Sensory Mechanisms of the Retina*, which was published in 1947, or his *Receptors and Sensory Perception*, which came out in 1955. Extensive information has been gained during the 1960s on the electrophysiology of color vision at the retina and most notably at the lateral geniculate. None of this excellent work is presented. Similarly, only token consideration is given to the electrophysiology of spatial vision. Consequently, it is not surprising that the large sections devoted to spatial vision and chromatic vision make almost no attempt to combine psychophysics with more recent electrophysiology. In fact, the color-blindness discussion is organized around G. L. Walls' 1955 theory of color vision (A branched-pathway schema for the color vision system and some evidence for it, *American Journal of Ophthalmology* 39:8–23), which is in conflict with this electrophysiology.

Although I would not recommend this book as a text, I believe that its reading would greatly reward both the expert and the student of sensory psychology. Rubin has successfully adopted a conversational style that makes much of the book a delight to read. At his best, Rubin presents material in a one-way conversation that is a pleasure to 'hear' and easy to understand. He provides lucid coverage of some basic material such as the measurement of light and the nature of color perception. More important, one finds scattered throughout the book interesting information that is difficult, or impossible, to find summarized elsewhere. Examples include the discussion of entoptic visual phenomena as well as short discussions on the history of such problems as rhodopsin and the scotopic luminosity function, the limits of visual resolution, color blindness, and binocular fusion. Finally, it is enlightening for a sensory psychologist to hear an ophthalmologist's approach to such topics as Snellen acuity and tests of color blindness.

To summarize, I would recommend *Fundamentals of Visual Science* to both the student of sensory psychology and the expert. But it strikes me as unlikely that any psychologists would want to use this book as either a research reference or as a text, although it is quite conceivable that it would find application both as a reference for lectures and as a source of selected assigned readings.

Donald C. Hood, *Columbia University*

Orthogonal Localization of Visual Objects: Some Experiments on Space Perception

By Jaakko Jarvinen. Helsinki: Suomalainen Tiedeakatemia, 1959. Pp. 141.

This dissertation consists of a report of three experiments on the misjudgment of orientation, covering textured surfaces tilted in the sagittal plane, the depth relationship of vertical rods as influenced by the plane of the background, and the component lines in parallelograms and zig-zag figures. The data are predictable, the theoretical analysis superficial, and tedious details of conventional technical and statistical methodology make up the bulk of the book.

W. B. Templeton, *York University, Canada*

Listening and Attention

By Neville Moray. Baltimore: Penguin Books, 1969. Pp. 107. Paperback, $1.45.

Listening and Attention is an agreeably no-nonsense book in a field where nonsense ordinarily thrives. Neville Moray knows the standard theories of selective listening inside and out—Cherry, Broadbent, Treisman, and Deutsch are all given their due in lucid exposition—but he correctly sees that none of those theories is very well developed or goes far toward predicting any new findings. To point this up clearly, Moray devotes most of a brief third chapter (the first two review the basic phenomena) to certain experiments of his own, in which subjects were required to tap a key whenever a signal occurred on *either* of two channels. These experiments are not yet in the open literature, and it is useful to have an account of them here, as they introduce a new set of problems for theories of attention. Moray himself leans toward a rapid-alternation theory, but his advocacy of it is gentle and does not coerce the reader.

Other chapters deal with the split-span experiment, the unsuccessful attempts to measure attention-switching time, and the even less successful efforts to study the physiology of attention; there is also a conclusion. Throughout the book Moray proves himself a lucid writer and a soft-spoken but trenchant critic.

The reader must be careful not to confuse this slim paperback, intended for and useful in undergraduate courses, with Moray's more serious effort to put forth his views in the 218-page hard-bound volume entitled *Attention: Selective Processes in Vision and Hearing* (New York: Academic Press, 1970). The two books were written more or less simultaneously (as is fitting for works in this area), but they are quite different. The one reviewed here is a modest but useful addition to Penguin's Science of Behavior series.

Ulric Neisser, *Cornell University*

Attention and Performance III

Edited by A. F. Sanders. Amsterdam: North Holland Publishing, 1970. Pp. 442. $19.50.

If this volume reads more like an issue of a journal than a monograph volume, it may be because it is. It is a third-generation model of a collection of papers prepared for a symposium dedicated to human-performance theory, a symposium which had its first coming out in 1966, with the papers subsequently published as volume 27 of *Acta Psychologica* (1967) and, simultaneously, *Attention and Performance* (A. F. Sanders, ed., Amsterdam: North Holland Publishing, 1967). Volumes II and III of the series followed. Though neither attention nor performance is ever vigorously defined, the stated need or raison d'etre for the original volume was that "the typical feature of performance research is that it does not fit one of the classical headings like perception, learning, motor behavior and motivation, but integrates aspects of each subject" (*Acta Psychologica* 27:vii). That position is reasserted for the volume now under review.

The attraction to the construct of attention is understandable. Rooted in a long history, applicable to almost all specialties, compatible with existing techniques and paradigms, and often suggesting more than is in fact delivered, attention is at once an integrative and uniting concept. Advocates rallying around the theme discover a common (generalized?) interest, even though they come from strikingly different orientations. Life is breathed into probability matching and discrimination learning, while psychophysics and reaction time take on a new dynamic. There is probably much value to a rethinking of traditional activities—and even greater worth in revisiting the non-learning activities that are studied more vigorously by those scholars outside the United States, for whom *Acta Psychologica* is the intended vehicle of communication (21 of the 30 contributors are non-Americans, though many are familiar to the American readership).

One may quibble with the choice of labels, but the variety of substantive theory and technique brought to bear on selected issues is both refreshing and sometimes valuable. Whereas the original volume of the series arranged papers in six topics—single-channel theory and information processing, reaction processes, physiological correlates of attention and reaction time, short-term memory and information processing, eye movements and visual search, long-term performance (i.e., vigilance)—the second volume focused on reaction processes. The present volume, *Attention and Performance III*, is organized into five topics—reaction processes, selective attention, subjective probability and performance, memory mechanisms and information processing, sustained attention—and in the main seems to be a retake of the earlier version. To be sure, the contributions are different and some new issues are discussed; but the contents, style, and nature of the writings more resemble another issue of *Acta Psychologica* than a new treatise on attention and performance. Substantively, the attentional value of the contributions is uneven, and one would have hoped to see reports by other well-qualified investigators who deserve representation in the volume. Technically, the index has several errors and could have been made more useful, but the work is otherwise carefully produced.

It is convenient to have such an ensemble of papers, particularly the reviews, theoretical analyses, and integrated programmatic studies. Perhaps there is a need for a regular serial volume, something like an 'Advances in . . .' Perhaps too, it might then be possible to leave the journal articles for *Acta Psychologica*, to invite more chapters of un-rehashed material, and to bring the volume's price down to such a level that the book could be reasonably expected to reach a wider audience.

David I. Mostofsky, *Boston University*

The Bisected Brain

By Michael S. Gazzaniga. New York: Appleton-Century-Crofts, 1970. Pp. xii, 172. $10.00.

The sheer conceptual beauty of the split-brain procedure lies not only in its power in localizing the site at which memories are stored but in the fact that the same surgical and experimental methods can be used on men as well as animals—and with complete ethical justification. For it turns out that certain epileptics with a focus in one hemisphere are sometimes aided by division of the huge commissures that connect their cerebral hemispheres together. This operation limits the spread of seizures to one half of the brain and in many cases, for an unknown reason, reduces their frequency and severity as well.

The technique was first used on humans more than 30 years ago, but, perhaps because of incomplete transection or inappropriate tests, no severe psychological deficits were observed. But in recent years a new series of patients have been treated and R. W. Sperry's group at Pasadena has tackled the question of whether the patients' minds had been split as well as their brains. M. S. Gazzaniga was there at the start and he has been in the front line of experiments on lower animals as well as split-brain humans ever since. In this short and snappy book, he provides a much-needed review of the role of the commissures in the transfer of information between the hemispheres. Unfortunately, like so many first editions nowadays, it is badly marred by printer's and proofreader's errors.

The book has a self-confessed bias not only toward the questions that interest Dr. Gazzaniga but also toward his particular prejudices and hypotheses. He begins in a novel way by describing the actual surgical methods for dividing the hemispheres of monkeys, cats, and pigeons. Unfortunately, this admirable attempt is not helped much by the rather confusing diagrams.

Gazzaniga argues forcefully that commissurotomy causes total physiological disconnection of the two sides of the brain. He produces convincing evidence that apparent communication between the hemispheres is often due to what he calls "cross-cuing." The ability to reach for an object (seen by one hemisphere) with the ipsilateral hand (controlled by the other hemisphere) is due to the fact that the seeing hemisphere can turn the head, neck, and shoulders toward the object and the hand can just be shot out blindly to pick it up. The position of a flash of light shown to the left visual hemifield (projecting to the right hemisphere) can be described verbally. Gazzaniga denies that this is due to the right hemisphere's controlling speech or to subcortical interhemispheric connections: the left hemisphere merely takes note of where the right hemisphere directs the head and eyes. The apparent lack of disability in humans with agenesis of the corpus callosum is also explained in these terms, as well as a host of other strange findings. In general, this argument is ingenious and feasible. However, there seem to be one or two exceptional cases that cannot be accounted for by "cross-cuing"; indeed there is considerable evidence for ipsilateral sensory and motor pathways and subcortical sensory function, at least in lower animals. One rather blatant omission is reference to Trevarthen's argument that collaboration between the split hemispheres can only be optimized by simultaneous presentation to the two hemispheres of the objects to be discriminated.

Gazzaniga also reviews what is known of the neurophysiology of the corpus callosum. Here is a big challenge for future research, because no one has studied neurons in the anterior part of the commissure and even the visual fibers in the splenium have receptive fields no more than very close to the midline of the visual field. So there is no explanation for the ability of a normal human to describe an object appearing in the periphery of the left hemifield. As Gazzaniga points out, breaking the code that callosum uses would put us a long way toward understanding the brain.

Finally, Gazzaniga allows himself a little speculation about the origin of cerebral dominance. He suggests that the newborn infant has a functionally split brain. The slightest initial tendency for the child to use one hand more than the other would lead, by a kind of positive feedback, to explosive development of the corresponding hemisphere, with subsequent superiority in language and motor control. This idea is fascinating and fits in with many neurological findings, but it still does not explain the initial dominance of the left hemisphere nor the superior performance of the right hemisphere in spatial perception.

In conclusion, this is an important book, coming at an opportune time. It is highly individual but has an adequate bibliography to direct the reader further. It is a useful introduction, for researchers and students alike, to this remarkable neurological condition and its use as a powerful research tool.

Colin Blakemore, *The Physiological Laboratory, Cambridge, England*

Brain and Behavior: Vol. 1, Mood, States, and Mind

Edited by Karl H. Pribram. Baltimore: Penguin Books, 1969. Pp. 496. $2.95.

This is the first volume of a four-volume series. It is divided into four sections, the first of which is comprised of six philosophical-theoretical essays intended as a general introduction to the whole series. The remaining three sections cover (*a*) "Sleep, Alertness, and Mood," (*b*) "Physiological Drive States," and (*c*) "Attention and the Control of Input." There are 28 articles in all; of these, 10 were originally published before 1950, the earliest being an excerpt from Gall and Spurzheim (1809). Of the 18 articles that appeared after 1950, 7 were authored or co-authored by Pribram himself. Other authors or co-authors who appear more than once are Karplus and Kreidl, two chapters; Bailey and Davis, two chapters; Lindsley, two chapters; and Dement, three chapters. One final descriptive statistic: the chapters range from 2 to

88 pages in length; the longest—nearly 20% of the book—is a review by Jouvet of the neurophysiological mechanisms mediating sleep.

In short, this is a rather idiosyncratic collection, heavily oriented toward Pribram's own views. This singleness of orientation, as well as the limited number of substantive topics covered, lends unity to the book even though the individual chapters vary considerably in depth and breadth of coverage. Moreover, Pribram is to be congratulated for reprinting a considerable wealth of hard-to-get historical material rather than succumbing to the temptation to include only the most recent—and quickly out-of-date—findings.

It would be wasteful to review the individual chapters that make up this volume. Two chapters, however, struck this reviewer as deserving special mention. The first is by the Soviet psychologist Luria, in which he argues effectively for the need to consider the social-historical determinants of psychological processes. Much of Soviet writing on philosophical psychology, with its frequent invocation of Marx and Lenin as final authorities, can easily discourage Western readers. (The converse is also undoubtedly true. Soviet psychologists must tire of reading the dogmatic pronouncements of 'one way' behaviorists, to use their rather pejorative term.) Western psychologists interested in physiological mechanisms, especially with regard to such 'primitive' states as mood and emotion, have tended to give only lip service to the role of social and historical factors. The real consideration of those factors is an aspect of Marxist philosophy that deserves to be taken more seriously by Western scientists.

The second chapter that deserves comment here is the concluding paper by Pribram, "Toward a Neuropsychological Theory of a Person." In it, Pribram briefly describes the efferent control of sensory input, which supposedly leads to either enhanced or reduced redundancy along afferent channels. Enhanced versus reduced redundancy is, in turn, related to such global psychological dimensions as ethical versus aesthetic propensities, introversion versus extraversion, assimilation versus accommodation, and inner-directedness versus outer-directedness, among others. One could never accuse Pribram of lacking imagination (suffering from too much redundancy?) but it is not unfair to point out that the editor of this volume—as he ventures into the thicket of brain-behavior relationships—is sometimes an emperor-explorer without clothes.

James R. Averill, *University of California, Berkeley*

Brain and Behavior: Vol. 2, Perception and Action

Edited by Karl H. Pribram. Baltimore: Penguin Books, 1969. Pp. 575. $2.95.

This is a volume of 27 readings on the central brain processes concerned with perception, memory, cognition, and motor functions. Each section has a very brief introduction.

Only 12 of the 27 papers were written since 1960 and only one since 1966. There are surely more informative and up-to-date papers on receptive fields than the one chosen, which was written in 1966 and which would certainly not give a student an understandable view of this topic. The 1952 paper by Weiss on peripheral-central coordination is included, but there is no selection from more recent work—by Székely, for instance—work which is transforming ideas in this field. None of the papers on motor functions is more recent than 1956. The one paper on auditory functions contains no reference to audition.

The volume does not succeed in giving an up-to-date coverage or sampling of the area, but for the price it is a useful collection of several classic papers and some recent material.

Ian P. Howard, *York University, Canada*

Books received

(The books listed here have not as yet been noted in our pages. Listing here does not preclude their later review, however.)

Antrobus, John S., ed. *Cognition and Affect*. Boston: Little, Brown, and Co., 1970. Pp. x, 210. $5.95; paperback, $3.75.

Arnold, William J., and Levine, David, eds. *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1970. Pp. 334. $6.95; paperback, $3.25.

Aronson, R. R.; Tobach, E.; Lehrman, D. S.; and Rosenblatt, J. D. *Development and Evolution of Behavior: Essays in Memory of T. C. Schneirla*. San Francisco: W. H. Freeman, 1970. Pp. 656. $12.00.

Atkinson, J. W., and Birch, D. *The Dynamics of Action*. New York: Wiley, 1970. Pp. 380. $11.95.

Bales, Robert Freed. *Personality and Interpersonal Behavior*. New York: Holt, Rinehart and Winston, 1970. Pp. xii, 561. $11.95.

Barnouw, Erik. *The Image Empire: A History of Broadcasting in the United States*. Vol. 3, *From 1953*. New York: Oxford University Press, 1970. Pp. 396. $9.75.

Bartlett, Roland W. *The Success of Modern Private Enterprise*. Danville, Ill.: Interstate, 1970. Pp. 281. $5.95.

Baumeister, Alfred A., and Butterfield, Earl. *Residential Facilities for the Mentally Retarded*. Chicago: Aldine, 1970. Pp. 405. $12.95.

Beauchamp, Kenneth L.; Bruce, Richard L.; and Matheson, Douglas W. *Current Topics in Experimental Psychology*. New York: Holt, Rinehart and Winston, 1970. Pp. 291. Paperback, $6.95.

Bejerot, Nils. *Addiction and Society*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xxii, 299. $9.75.

Bell, R. G. *Escape from Addiction*. New York: McGraw-Hill, 1970. Pp. 201. $5.95.

Belliveau, Fred, and Richter, Lin. *Understanding Human Sexual Inadequacy*. New York: Bantam Books, 1970. Pp. 242. Paperback, $1.25.

Bennet, E. A. *What Jung Really Said*. New York: Schocken Books, 1970. Pp. 188. Paperback, $1.95.

Berg, Frederick S., and Fletcher, Samuel G., eds. *The Hard of Hearing Child: Clinical and Educational Management*. New York: Grune & Stratton, 1970. Pp. ix, 363. $12.75.

Best, Carole, and Kirschenbaum, Jack. *Instructor's Manual for Introductory Psychology*. Beverly Hills, Calif.: Cognitive Systems, 1970. Pp. ii, 7, 24.

Biener, K. *Wirksamkeit der Gesundheitserziehung*. Basel: S. Karger, 1970. Pp. 69.

Blachly, Paul H., ed. *Drug Abuse: Data and Debate*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xvi, 322. $12.50.

Bleuler, Eugene. *Autistic: Undisciplined Thinking in Medicine and How to Overcome It*. 5th German ed., trans. and ed. Ernest Harms. Darien, Conn.: Hafner, 1970. Pp. ix, 218. $11.50.

Bloom, Lois. *Language Development: Form and Function in Emerging Grammars*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. xiv, 270. $8.95.

Boisefeuillet, Jones, ed. *The Health of Americans*. Englewood Cliffs, N.J.: Prentice-Hall, 1970. Pp. vii, 209. $5.95; paperback, $2.45.

Briggs, Jean L. *Never in Anger: Portrait of an Eskimo Family*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 379. $15.25.

Brix, V. H. *You Are a Computer*. New York: Emerson Books, 1970. Pp. 144. $4.95.

Burns, R. C., and Kaufman, S. H. *Kinetic Family Drawings (K-F-D)*. New York: Brunner/Mazel, 1970. Pp. 160. $8.95.

Carter, Darrell B., ed. *Interdisciplinary Approaches to Learning Disorders: A Symposium by Thirteen Contributors*. Philadelphia: Chilton, 1970. Pp. xii, 191. $8.95.

Chess, Stella, and Thomas, A. *Annual Progress in Child Psychiatry and Child Development*. New York: Brunner/Mazel, 1970. Pp. 585. $15.00.

Clinebell, Howard J., and Clinebell, Charlotte H. *The Intimate Marriage*. New York: Harper & Row, 1970. Pp. xviii, 231. $5.95.

Cohen, Jozef. *Sensation and Perception;* Vol. 1, *Vision*, Vol. 2, *Audition and the Minor Senses. Secondary Motivation;* Vol. 1, *Personal Motives*, Vol. 2, *Social Motives. Operant Behavior and Operant Conditioning. Complex Learning. Personality Dynamics. Personality Assessment*. The Eyewitness Series in Psychology, nos. SS4, SS5, SS7, SS8, SS11, SS12a, SS17, SS18. Chicago: Rand McNally, 1970. $1.25 each.

Coles, Robert; Brenner, Joseph H.; and Meagher, Dermot; eds. *Drugs and Youth: Medical, Psychiatric, and Legal Facts*. New York: Liveright, 1970. $5.95.

Cooper, David. *The Death of the Family*. New York: Random House, Pantheon Books, 1970. Pp. 141. $5.95.

Cottle, William C., and Downie, E. M. *Preparation for Counseling*. 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1970. Pp. 420. $8.95.

Cowan, J. L. *Studies in Thought and Language*. Tucson: University of Arizona Press, 1970. Pp. 228. $6.50.

Craik, Kenneth H.; Kleinmuntz, Benjamin; Rosnow, Ralph L.; Rosenthal, Robert; Cheyne, J. A.; and Walters, R. H. *New Directions in Psychol-*

ogy IV. New York: Holt, Rinehart and Winston, 1970. Pp. viii, 382. Paperback, $3.95.

Crovitz, Herbert F. *Galton's Walk: The Analysis of Thinking, Intelligence, and Creativity*. New York: Harper & Row, 1970. Pp. xii, 159. $2.95.

DeBuskey, Matthew, ed. *The Chronically Ill Child and His Family*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xvi, 203. $11.50.

DeGreene, Kenyon B., ed. *Systems Psychology*. New York: McGraw-Hill, 1970. Pp. xiii, 593. $18.50.

Deutsch, C. P., and Schumer, F. *Brain-Damaged Children*. New York: Brunner/Mazel, 1970. Pp. 162. $7.95.

Dodwell, P. C., ed. *Perceptual Learning and Adaptation*. Baltimore: Penguin Books, 1970. Pp. 490. Paperback, $2.65.

Dodwell, P. C. *Visual Pattern Recognition*. New York: Holt, Rinehart and Winston, 1970. Pp. 276. $9.00.

Doob, Leonard W., ed. *Resolving Conflict in Africa: The Fermeda Workshop*. New Haven, Conn.: Yale University Press, 1970. Pp. 209. $7.50.

Duffy, John C. *Emotional Issues in the Lives of Physicians*. Springfield, Ill.: Charles C Thomas, 1970. Pp. vii, 92. $6.50.

Easson, William M. *The Dying Child: The Management of the Child or Adolescent Who Is Dying*. Springfield, Ill.: Charles C Thomas, 1970. Pp. 103. $5.75.

Edge, David, ed. *The Formative Years: How Children Become Members of Their Society*. New York: Schocken Books, 1970. Pp. xvi, 72. $4.50; paperback, $1.75.

Estes, William K. *Learning Theory and Mental Development*. New York: Academic Press, 1970. Pp. x, 233. $9.50.

Evans, Ellis D., ed. *Adolescents: Readings in Behavior and Development*. Hinsdale, Ill.: Dryden Press, 1970. Pp. xi, 525.

Eysenck, H. J. *The Structure of Human Personality*. New York: Barnes & Noble, 1970. Pp. xviii, 476. $9.50; paperback, $5.00.

Faas, Larry A., ed. *The Emotionally Disturbed Child: A Book of Readings*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xiv, 386. $14.50.

Feng, Gia-fu, and Kirk, Jerome. *Tai Chi: A Way of Centering, and "I Ching."* New York: Macmillan, 1970. Pp. 157. $6.95; paperback, $3.95.

Foley, John M.; Lockhart, Russell A.; and Messick, David M.; eds. *Contemporary Readings in Psychology*. New York: Harper & Row, 1970. Pp. ix, 470. Paperback, $4.95.

Fraisse, Paul, ed. *L'Annee psychologique*. Paris: Presses Universitaires de France, 1970. Pp. 339.

Francis-Williams, Jessie. *Children with Specific Learning Difficulties*. Elmsford, N.Y.: Pergamon Press, 1970. $8.00; paperback, $5.50.

Freeman, Howard E., and Sherwood, Clarence C. *Social Research and Social Policy*. Englewood Cliffs, N.J.: Prentice-Hall, 1970. Pp. 159. $5.95; paperback, $2.50.

Fried, Edrita. *Active/Passive: The Crucial Psychological Dimension*. New York: Grune & Stratton, 1970. Pp. ix, 222. $7.95.

Gagne, Robert M. *The Conditions of Learning*, 2d ed. New York: Holt, Rinehart and Winston, 1970. Pp. viii, 407. $7.95.

Gil, David G. *Violence Against Children: Physical Child Abuse in the U.S.* Cambridge, Mass.: Harvard University Press, 1970. Pp. xiii, 205. $6.50.

Gitman, L., and Woodfort-Williams, E., eds. *Interdisciplinary Topics in Gerontology.* Vol. 5, *Research, Training, and Practice in Clinical Medicine of Aging.* Basel: S. Karger, 1970. Pp. 94. $6.95.

Gleitman, Lila R., and Gleitman, Henry. *Phrase and Paraphrase: Innovative Uses of Language.* New York: W. W. Norton, 1971. Pp. xii, 242. $7.95.

Glueck, Sheldon, and Glueck, Eleanor. *Toward a Typology of Juvenile Offenders.* New York: Grune & Stratton, 1970. Pp. xvi, 203. $8.75.

Goda, Sidney. *Articulation Therapy and Consonant Drill Book.* New York: Grune & Stratton, 1970. Pp. xi, 102. $6.75.

Gould, Wesley L., and Barkun, Michael. *International Law and the Social Sciences.* Princeton, N.J.: Princeton University Press, 1970. Pp. xx, 338. $9.75.

Greene, Bernard L. *A Clinical Approach to Marital Problems.* Springfield, Ill.: Charles C Thomas, 1970. Pp. xiv, 445. $20.00.

Grossack, Martin, and Gardner, Howard. *Man and Men: Social Psychology as a Social Science.* Scranton: International Textbook, 1970. Pp. 363. $3.95.

Group for the Advancement of Psychiatry, Committee on Psychiatry and Law. *The Right to Abortion: A Psychiatric View.* New York: Scribner's, 1970. Pp. 75. $4.95; paperback, $1.95.

Hamilton, Eleanor. *What Made Me?* New York: Hawthorn Books, 1970. $4.95.

Hammond, K. R.; Householder, J. E.; and Castellan, N. J., Jr. *Introduction to the Statistical Method.* 2d ed. New York: Knopf, 1970. Pp. 410.

Hansen, David A., and Kolbmann, John J. *Closed Circuit Television for Police.* Springfield, Ill.: Charles C Thomas, 1970. Pp. x, 103. $8.75.

Hare, Robert D. *Psychopathy: Theory and Research.* New York: Wiley, 1970. Pp. x, 138. $5.00; paperback, $2.95.

Hayes, J. R., ed. *Cognition and the Development of Language.* New York: Wiley, 1970. Pp. 370. $12.00.

Heber, Rick. *Epidemiology of Mental Retardation.* Springfield, Ill.: Charles C Thomas, 1970. Pp. 119. $8.75.

Heine, Ralph W. *Psychotherapy.* Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. x, 170. $6.95; paperback, $2.75.

Hermelin, B., and O'Connor, N. *Psychological Experiments with Autistic Children.* New York: Pergamon Press, 1970. Pp. vi, 141. $8.00.

Horn, Gabriel, and Hinde, Robert A., eds. *Short-term Changes in Neural Activity and Behaviour.* New York: Cambridge University Press, 1970. Pp. vi, 628. $28.50.

Horowitz, Irving Louis, and Friedland, William H. *The Knowledge of Factory: Student Power and Academic Politics in America.* Chicago, Ill.: Aldine, 1970. Pp. viii, 351. $9.75.

Hudson, Liam. *The Ecology of Human Intelligence.* Baltimore: Penguin Books, 1970. Pp. 368. $2.45.

Hutschnecker, Arnold A. *The Will to Happiness.* Rev. ed. New York: Trident Press, 1970. Pp. 192. $4.95.

continued on page 350

ANNOUNCEMENT

First International Conference on Gambling

Gambling, with its increasing popularity and legalization, deserves a thorough examination. The moral, social, political, economic, mathematical, legal, practical, and other aspects of gambling will be considered at this First International Conference on Gambling.

Objectives of the conference are to involve the general public in the study of gambling—registrations will be open to all; to study the status of gambling in various countries throughout the world; and to add to our limited knowledge of gambling by examining research findings and by the exchange of ideas among international scholars, writers, and authorities on gambling.

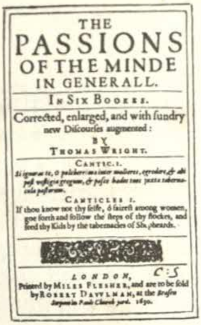
Areas of discussion will include gambling as a historical perspective and the current state of the phenomenon; sociological studies of the topic; cross-cultural variations in gambling practices; pathological gambling; the psychology of gambling and risk taking; gambling odds and probability; gambling and crime; the problems of legalization of gambling; and the discussion of specific forms of gambling.

Location:	York University, Toronto, Ontario, Canada
Date:	June 25–29, 1972
Fee:	$125.00 for full conference, residence, and admission to Queen's Plate
Accommodation:	On campus accommodation is available for conference participants and their spouses
Programme:	Presentation of papers, speeches, informal discussions, demonstrations, film, exhibits, etc. Visits to races and a Monte Carlo night.

For further information on the First International Conference on Gambling, write to:

The Centre for Continuing Education,
York University,
4700 Keele Street,
Downsview, Ontario, Canada.

or phone: (416) 635-2505.



THE
PASSIONS
OF THE MINDE
IN GENERALL.
IN SIX BOOKES.
Corrected, enlarged, and with ſundry
new Diſcourſes augmented:
BY
THOMAS WRIGHT.
CANTIC. 1.
Si ignoras te, o pulcherrima inter mulieres, egredere, & abi post vestigia gregum, & paſce hædos tuos juxta tabernacula paſtorum.
CANTICLES 1.
If thou know not thy ſelfe, ô faireſt among women, goe forth and follow the ſteps of thy flockes, and feed thy Kids by the tabernacles of Shepheards.

LONDON, C:S
Printed by MILES FLESHER, and are to be ſold
by ROBERT DAWLMAN, at the *Braſen
Serpent in Pauls Church yard.* 1630.

THE PASSIONS OF THE MINDE IN GENERALL

THOMAS WRIGHT

A Reprint Based on the 1604 Edition with an Introduction by Thomas O. Sloan

In 1604 Thomas Wright, Catholic priest bent on proving that loyalty to the Church of Rome should not preclude loyalty to the English crown, added to his controversial writings the observation that "Italians and Spaniardes, with other inhabitants beyond the Alpes," found Englishmen and their neighbors "simple, uncircumspect, unwarie, easie to be deceived," . . . partly because they had been brought up "with too much feare and terrour. . . . The Italians and Spaniards contrariwise, by bringing up their children with more libertie [make them] at sixteen . . . as bold and audacious as ours at thirtie."

Wherefore, Wright "thought good to trie if a little direction would help our Countriemen . . . to discover other men's passions, and [learn] how to behave ourselves when such affections extraordinarily possesse us."

Wright's doctrine that men are meant to use their passions follows St. Thomas Aquinas, but Wright carried it further and developed the theory that the passions and, subsequently, the will may be moved and persuaded by rhetoric, which in the seventeenth century encompassed all the language arts.

After providing a brief biography of Wright, Sloan's introduction points out that the importance of Wright's book to students of literature rests in its clear reflection of the cross-currents of political, religious, psychological, and rhetorical movements of its time. 440 pages. $12.95.

UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

ANNOUNCEMENT

A Clearinghouse for Research on Civil Rights

The Center for National Policy Review is devoted to assuring that Federal laws and policies in urban areas are implemented to protect the interests of minority groups and the poor. We are a group of lawyers and social scientists engaged in research and legal action on behalf of the principal civil rights groups in the country, including the NAACP, the Urban League, and the National Urban Coalition.

As part of our purpose the Center will serve as a clearinghouse for social science research which has implications for Federal policies on equality of opportunity. We are soliciting brief summaries of such research which is either completed or in progress. Currently our principal concerns are with housing, credit practices, federal site selection, employment, and metropolitan school desegregation, but we welcome reports on research in other areas of civil rights concern such as welfare program administration, political participation, and revenue sharing. These research reviews will appear in a quarterly newsletter which will be sent to organizations involved in civil rights reform, urban studies centers, social scientists, and civil rights lawyers. The Center for National Policy Review is located at the School of Law, Catholic University of America, Washington, D.C. 20017, (202) 832-8525.

Samuel Gubins
Director of Research

Winner of 22nd National Book Award

Science in the British Colonies of America

RAYMOND PHINEAS STEARNS

Stearns discusses virtually every scientist of any consequence who lived or worked in colonial America. . . . truly a mine of information . . . will be useful for years to come. — George H. Daniels, Science.

. . . a graceful work of scholarship filled with an infinite variety of interesting details about a period and a place rarely considered in the history of science. — Christopher Lehmann-Haupt, The New York Times.

. . . a magnificent document. There is nothing like it in all the literature on American history or on the history of science in America, nor is there likely to be for some time. — I. Bernard Cohen, Harvard University.

$20.00

from your bookstore or from
UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

ANNOUNCEMENT

Visiting Lecturer Program in Statistics

A Visiting Lecturer Program in Statistics has been organized for the ninth successive year. The program is sponsored jointly by the principal statistical organizations in the United States, the American Statistical Association, the Biometric Society and the Institute of Mathematical Statistics. The National Science Foundation provides partial financial support. Leading teachers and research workers in statistics—from universities, industry and government—have agreed to participate as lecturers. Lecture topics include subjects in experimental and theoretical statistics, as well as in such related areas as probability theory, information theory and stochastic models in the physical, biological, and social sciences.

The purpose of the program is to provide information to students and college faculty about the nature and scope of modern statistics, and to provide advice about careers, graduate study, and college curricula in statistics. Inquiries should be addressed to: Visiting Lecturer Program in Statistics, Department of Statistics, Southern Methodist University, Dallas, Texas 75222.

The list of participating lecturers includes: Z. W. Birnbaum (University of Washington), Melvin W. Carter (Brigham Young University), Arthur Cohen (Rutgers University), Thomas M. Cover (Massachusetts Institute of Technology), Edwin L. Crow (U.S. Department of Commerce Research Laboratories, Boulder, Colorado), Herbert A. David (University of North Carolina), H. T. David (Iowa State University), W. A. Ericson (University of Michigan), Donald A. Gardiner (Oak Ridge National Laboratory), Joseph L. Gastwirth (Office of Statistical Policy, Washington, D.C.), John P. Gilbert (Harvard University), Leon J. Gleser (Johns Hopkins University), R. Gnanadesikan (Bell Telephone Laboratories, Murray Hill, New Jersey), William C. Guenther (University of Wyoming), H. Leon Harter (ARL, Wright-Patterson AFB, Ohio), Leon H. Herbach (New York University), Charles R. Hicks (Purdue University), Warren M. Hirsch (New York University), Ronald R. Hocking (University of Houston), Myles Hollander (Florida State University), Peter W. M. John (University of Texas at Austin), Leo Katz (Michigan State University), S. Kotz (Temple University), Clyde Y. Kramer (Virginia Polytechnic Institute), Irving H. LaValle (Tulane University), Eugene Lukacs (The Catholic University of America), B. J. Mandel (Washington, D.C.), Carl E. Marshall (Stillwater, Oklahoma), John Neter (University of Minnesota), P. L. Odell (Texas Tech University), S. James Press (University of Chicago), Tim Robertson (University of Iowa), Richard N. Schmidt (SUNY at Buffalo), Paul Switzer (Stanford University), David L. Sylwester (University of Washington), Benjamin J. Tepping (Washington, D.C.), W. A. Thompson, Jr. (University of Missouri), B. J. Trawinski (University of Alabama at Birmingham), Bruce E. Trumbo (California State College at Hayward), Grace Wahba (University of Wisconsin), K. T. Wallenius (Clemson University), John W. Wilkinson (Rensselaer Polytechnic Institute), James Yackel (Purdue University).

The organizing committee consists of: H. T. David, S. W. Greenhouse, S. S. Gupta, W. J. Hall, L. Katz, L. H. Koopmans, I. Olkin, D. B. Owen (Chairman), G. J. Resnikoff.

A Scrupulous Meanness

A Study of Joyce's Early Work

EDWARD BRANDABUR

Brandabur combines literary criticism and psychoanalytic perceptions in his analysis of Joyce's early work, particularly *Dubliners* and the only extant play, *Exiles*.

Joyce felt that there was "something gone wrong" in Ireland; there was a paralysis at its center. Dubliners in particular were unconscious and eager victims of "sadomasochistic neurosis," tending to live in vicarious experience of triumph and abasement. Brandabur shows that Joyce's style of "scrupulous meanness" enabled him to control his chiefly negative attitude toward Dublin. The author holds that a moral purpose of Joyce's writing was to free his compatriots (and perhaps himself) from their enslavement to sadomasochism.

". . . Brandabur has dug deep and marshalled his evidence with vigor. . . . the best psychoanalytic reading of Joyce's early work . . . (has) a unique perspective that every student and critic of Joyce will have to reckon with." — Robert Scholes, Brown University. $6.95.

UNIVERSITY OF ILLINOIS PRESS **Urbana · Chicago · London**

Ake & issu
Docket to the Librarian
R
22/8

565/2618

Vol. 84, No. 4 Founded in 1887 by G. Stanley Hall December 1971

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys
Charles W. Eriksen

M. E. Bitterman
Benton J. Underwood

LEO POSTMAN and KAREN STARK. Encoding Variability and Transfer 461

MICHAEL M. GRUNEBERG and ROBERT N. SYKES. 'Coding' in Studies of Acoustic and Semantic Interference and Confusion 473

ROBERT SEKULER and ALBERT ERLEBACHER. The Two Illusions of Müller-Lyer: Confusion Theory Reexamined 477

STUART APPELLE. Visual and Haptic Angle Perception in the Matching Task 487

CHARLES H. WEBB and DONALD H. McBURNEY. Salivary Habituation: Quantitative Similarities to Sensory Adaptation 501

STANLEY A. SUMMERS and JAMES S. FLEMING. Construction and Reconstruction in Memory 513

continued on back cover

Editors **Lloyd G. Humphreys** **M. E. Bitterman**
Charles W. Eriksen **Benton J. Underwood**
Editor Emeritus **Karl M. Dallenbach**
Book Review Editor **Julian Hochberg**
Consulting Editors **William Epstein** **Robert Fox** **Wendell Garner**
Lawrence E. Jones **John Krauskopf** **Allan Wagner**
Index Editor **Rand B. Evans** Necrology Editor **E. R. Hilgard**

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Eriksen (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Hochberg (Columbia University, New York, New York 10027).

Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text. Authors of articles may order offprints at the time they return galley proof; other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

The subscription price is $10.00 a year (foreign $10.50); single copies cost $2.75. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

© 1971 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

26 AUG 1972

Vol. 84, No. 4

continued from front cover

December 1971

JOHN F. HALL. Formal Intralist Response Similarity: Its Role in Paired-Associate Learning 521

MICHEL LOEB and JOHN R. BINFORD. Modality, Difficulty, and 'Coupling' in Vigilance Behavior 529

NORMAN H. ANDERSON and DAVID J. WEISS. Test of a Multiplying Model for Estimated Area of Rectangles 543

W. T. WOODARD. Classical Respiratory Conditioning in the Fish: CS Intensity 549

DONALD J. LEVIS. CS Complexity and Intertrial Responding in Shuttlebox Avoidance Conditioning 555

STANLEY COREN. A Size-Contrast Illusion without Physical Size Differences 565

IN-MAO LIU and CHOR-PING HUI. An Instrumental-Learning Task for Human Subjects 567

GALE BISHOP and JEROME SMITH. Dimensional Shifts with Verbal Units 581

Book Reviews (Stanley Coren, Philip Shaver, Erich Klinghammer, Ralph F. Hefferline, Walter Cohen; books received) 589

Index to Volume 84 599

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS

Urbana, Illinois 61801

American Journal of Psychology
1971, Vol. 84, No. 4

ENCODING VARIABILITY AND TRANSFER

Leo Postman and Karen Stark
University of California, Berkeley

Martin found negative transfer to increase as a function of stimulus meaningfulness in the A-B, A-Br paradigm relative to the A-B, C-B paradigm—a finding in line with the hypothesis of encoding variability. But when both stimulus and response meaningfulness in those paradigms were manipulated in the present experiment, there was no difference in relative negative transfer with either kind of meaningfulness. This result, an evaluation of the earlier finding, and other evidence lead to the conclusion that the hypothesis still lacks firm empirical support.

In a recent theoretical analysis of transfer, Martin (1968) considered the influence of stimulus meaningfulness on the acquisition and retention of successive paired-associate lists. The central feature of his analysis was the hypothesis of encoding variability—the assumption that there is variation in the perceptual encoding of the nominal stimuli in a list and that the amount of such variation is inversely related to the meaningfulness of those stimuli. In the acquisition of a single list, stability of stimulus encoding is essential to improvement; hence increases in stimulus meaningfulness may be expected to have a favorable effect. But in the acquisition of successive lists that share common stimulus terms, increases in stimulus meaningfulness may be expected to produce increasing amounts of negative transfer. That is, highly meaningful stimuli continue to be encoded in the same way as the subject moves from the first to the second list; consequently, the original associations must be unlearned in the transfer phase. By contrast, less meaningful stimuli are more likely to be recoded when a new task is introduced, so that interference from prior associations is minimized. In effect, then, for successive lists that share common stimulus terms the paradigms of negative transfer—A-B, A-D and A-B, A-Br—represent the actual learning conditions when stimulus meaningfulness is high; and when it is low, the functional paradigms shade into A-B, C-D and A-B, C-B respectively.

Martin (1968) reported a study of transfer whose results he viewed as consistent with these implications of the hypothesis of encoding variability. This experiment compared the A-B, C-B and A-B, A-Br paradigms of transfer, using lists of six CVC-digit pairs as learning materials. Stimulus meaningfulness was either high or low. A difference in favor of the former was observed in first-list learning. In the transfer phase the high-meaningfulness list was again learned faster than the low one under the control (C-B) condition; however, the scores of the two experimental (A-Br) groups were now identical. This means that a significant amount of relative negative transfer (A-Br versus C-B) was obtained when stimulus meaningfulness was high but not when it was low. This pattern of differences was in accord with theoretical expectations (although the critical interaction of paradigm and stimulus meaningfulness was not evaluated). The same trends were apparent in first-list relearning.

According to Martin's interpretation, differences in the amount of stimulus recoding had counteracted the usual effects of meaningfulness in transfer and relearning under the experimental condition. Another possibility may warrant consideration, however; namely, that first-list associations were a more effective source of interference when stimulus meaningfulness was high than when it was low. The differences in speed of first-list learning make it reasonable to suppose that associative strength at criterion was not equal for the high- and low-meaningfulness lists—and variations in the degree of first-list learning may be more likely to influence the amount of associative interference under the A-Br paradigm than under other conditions of transfer (Postman, 1962). Furthermore, in absolute terms the magnitude of Martin's paradigm × stimulus-meaningfulness interaction was not great; on trials to criterion on the second list, the difference between the A-Br and C-B groups was 3.95 for the high- and 1.95 for the low-meaningfulness conditions. First-list strength, albeit not a powerful determinant of transfer, might well produce differences of that order of magnitude.

The alternative interpretation just suggested implies that the interaction Martin observed should also be found when response rather than stimulus meaningfulness is varied. That is, as the level of response meaningfulness becomes higher, first-list associations should be acquired with increasing speed and become more effective sources of associative interference. There is some evidence in the literature that negative transfer does, indeed, increase with the level of response meaningfulness, both for the A-D (Goulet, 1965; Jung, 1963) and the A-Br paradigm (Merikle, 1968). In the latter experiment, comparisons between the C-B and the A-Br paradigms showed increases in relative negative transfer as a function of both stimulus and response meaningfulness. The conditions of that study differed, however, from Martin's in several important respects: pairs of trigrams were used as the learning materials, with stimulus and response meaningfulness varied factorially; the drop-out method was used in learning, and the transfer phase was preceded by a test of first-list recall. The results may, therefore, have limited generality.

The purpose of the present study was to reexamine the interaction between paradigm and stimulus meaningfulness that Martin described and to compare it with the interaction between paradigm and response meaningfulness under parallel conditions. The question of interest was whether the influence of meaningfulness on relative transfer is specific to the locus at which this variable is manipulated. Such specificity seems essential if the observed effects of stimulus meaningfulness are to be attributed to encoding variability rather than to processes that are a function of stimulus and response meaningfulness alike.

A reappraisal of the relation between stimulus meaningfulness and transfer is timely too because other recent evidence bearing on the hypothesis of encoding variability has been largely negative. First, when stimuli permit selection of alternative features, subjects have apparently failed to switch from one type of cue to the other so as to change the functional paradigm of transfer and to minimize the possibilities of associative interference (Goggin and Martin, 1970; Houston, 1967). Moreover, when stimulus selection is brought under experimental control and a change in functional cues is effected, transfer performance has remained unaffected (Williams and Underwood, 1970). Finally, in direct contradiction to Martin's hypothesis, negative transfer under the A-B, A-D paradigm has been found to be independent of stimulus meaningfulness (Weaver, McCann, and Wehr, 1970).

METHOD

—Design—The experiment was designed to assess the influence of both stimulus and response meaningfulness on transfer under the C-B and the A-Br paradigms. The learning materials were lists of paired associates composed of CVC trigrams and two-syllable adjectives. The design comprised eight groups representing the factorial combination of the two paradigms (*C-B* and *A-Br*), two types of pairs (*trigram-adjective* and *adjective-trigram*), and two levels of trigram meaningfulness (*high* and *low*).

—Lists—There were two sets of six CVC trigrams at each level of meaningfulness. The association values (Archer, 1960) of the high-meaningfulness items were 100%.[1] The values ranged from 19% to 24% in one of the low-meaningfulness lists, and from 19% to 23% in the other. Both intralist and interlist formal similarity were minimized. Ten different consonants and five different vowels were used within each list. There were only two duplications of consonants between the two sets of trigrams, and these two letters occupied different positions. Two sets of six two-syllable adjectives were used. There were no duplications of initial letters either within or between sets. Meaningful similarity both within sets and between sets was minimized by inspection.

At a given level of meaningfulness, there were four basic lists, representing the possible combinations of the two sets of trigrams and two sets of adjectives. For each basic list there were two different pairings of the stimulus and response terms. Thus there were altogether eight different lists at each level of meaningfulness, lists which were used equally often as the first and the second task.

—Procedure—Learning was by the anticipation method at a 2:2-sec rate, with a 4-sec intertrial interval. The lists were presented in six different random orders. The first list was learned to a criterion of one perfect recitation. The second list was learned to the same criterion or for 10 trials, whichever took longer. The interval between the two lists was 1 min 30 sec.

A test of recall for the first list was given 1 min 30 sec after the end of second-list learning. The stimulus terms were presented at a 10-sec rate, and the subject was required to recall the responses. No feedback was given during the recall trial. Since all subjects had at least 10 transfer trials, the degree of overlearning on the second list was not equated across conditions. Also, since the time between the end of interpolated learning and the test of recall was held constant, there was necessarily some variation in the length of the retention interval for the first list. The results of the test of recall must be considered in light of these limitations, although it seems unlikely that major sources of bias were introduced.

—Subjects—There were 16 subjects in each of the eight groups. They were undergraduates at the University of California and were not necessarily naive to verbal-learning experiments. Assignment to conditions was in blocks of eight, with one subject in each condition per block. The running order within blocks was determined by means of a table of random numbers.

—Comparison with Martin's procedure—The four treatments in which the trigrams occupied the stimulus positions in the paired-associate lists represented a replication of Martin's design. However, some changes were made in the present experiment. First, new sets of trigrams were constructed, because in the earlier study the lists at the two levels of meaningfulness were not fully equated with respect to either intralist or interlist formal similarity. Second, the trigrams were paired with adjectives rather than the digits 1–6. Third, conventional paired-associate practice was substituted for the selective-reinforcement procedure Martin used.

The reasons for the last two changes were these. In Martin's experiment the correct stimulus-response pairings were never presented. Instead, the subjects were instructed to guess a number between 1 and 6 to each of the stimulus terms, and a feedback light was used to confirm correct responses. This procedure has several potential disadvantages: (*a*) random variability in item strength at criterion is promoted by the fact that the point at which the correct response to a given stimulus is discovered is a matter of chance; (*b*) the guessing requirement probably inflates the error rate as compared to the conventional procedure; (*c*) guesses reflecting preexperimental associations are likely to be more persistent to high- than to low-meaningfulness stimuli. In view of these possible complications, it seemed preferable to revert to the conventional method of paired-associate practice. Furthermore, since the use of consecutive digits as responses may encourage the application of guessing strategies during the acquisition of a short list even under the conventional procedure, the trigrams were paired with adjectives rather than digits.

RESULTS

In the interest of facilitating comparison with Martin's findings, the various measures of performance are first considered for the trigram-adjective and the adjective-trigram lists separately, after which the differences between the two types of materials are evaluated.

First-list learning

The mean numbers of trials to criterion on the first list are shown in Table 1. There were no reliable differences between groups learning the same type of list. Criterion on the trigram-adjective lists was reached faster when stimulus meaningfulness was high than when it was low [$F(1, 62) = 5.76, p < .02$]. Such a difference was also reported by Martin.

Speed of acquisition was directly related to the level of meaningfulness for the adjective-trigram lists [$F(1, 62) = 6.92, p < .02$]. Learning was slower when the trigrams were in the response rather than the stimulus positions [$F(1, 124) = 10.21, p < .01$]. As usual, response meaningfulness had a greater effect on acquisition than stimulus meaningfulness, but the interaction of type of list with level of meaningfulness was not significant ($F < 1$).

Second-list learning

Two measures of transfer are presented in Table 1: the mean total numbers of correct responses in 10 trials of second-list learning, and the mean numbers of trials to criterion on that list. High stimulus meaningfulness continued to have a favorable influence on performance with the trigram-adjective lists in the transfer phase, but this effect was somewhat greater under the C-B than under the A-Br paradigm. Thus, both measures showed slightly larger amounts of relative negative transfer (A-Br versus C-B) for the high- than for the low-meaningfulness lists, in apparent agreement with the trend observed by Martin. It should be

Table 1. Measures of acquisition and recall

Condition	List 1		List 2	
	T/c	Recall	Correct, trials 1–10	T/c
Trigram-adjectives				
A-Br (H)	8.12	3.19	37.50	8.69
C-B (H)	6.62	5.12	49.88	4.75
A-Br (L)	10.12	3.25	34.38	9.94
C-B (L)	10.62	4.19	40.31	8.31
Adjective-trigrams				
A-Br (H)	9.38	3.19	37.69	9.94
C-B (H)	10.56	5.12	45.31	5.56
A-Br (L)	15.75	3.50	30.69	13.00
C-B (L)	14.12	4.56	42.69	7.88

Note: H and L designate high and low meaningfulness respectively; T/c is trials to criterion.

noted, however, that the sampling differences in first-list learning, although not reliable, showed a parallel pattern: a larger difference between the C-B than between the A-Br groups. In any event, the interaction of paradigm and stimulus meaningfulness was far from significant ($F = 1.59$ and 1.03 for correct responses and trials to criterion respectively). The main effects of stimulus meaningfulness and of paradigm were significant for both measures at the .05 level or better.

At this point, a direct comparison with the measures of transfer reported by Martin is in order. As noted previously, he did not test the interaction of paradigm and stimulus meaningfulness. Instead, the differences in the percentages of net transfer at each level of stimulus meaningfulness were evaluated separately. The net percentage of transfer from the first to the second list was determined by the formula $100 \times (T_1 - T_2) / (T_1 + T_2)$, where T is trials to criterion. Application of this formula presumably takes account of individual differences in learning ability (although the percentages calculated to widely varying bases are likely to be unstable). The mean percentages of transfer obtained in Martin's experiment and in the present study are juxtaposed in Table 2. Clearly, the earlier finding was not replicated: in our results the differences between the two paradigms were quite comparable at the two levels of stimulus meaningfulness.

We turn now to the transfer results for the adjective-trigram lists (Table 1). The high-meaningfulness lists had an apparent advantage over the low ones, but the difference fell short of significance [$F(1, 60) =$ 3.41 for total correct responses and 3.93 for trials to criterion, with a value of 4.00 required at the .05 level]. Performance was significantly better under the C-B than under the A-Br paradigm. The F ratios for correct responses and trials to criterion were 14.19 and 12.26 respectively ($ps < .01$). The interaction of paradigm and meaningfulness was again far from significant ($F < 1$ in each analysis). Thus, the amounts of relative negative transfer were independent of response meaningfulness. This conclusion is borne out by the mean percentages of transfer shown in Table 2.

Table 2. Mean percentages of transfer

Experiment	Trigram-adjectives				Adjective-trigrams			
	A-Br (H)	C-B (H)	A-Br (L)	C-B (L)	A-Br (H)	C-B (H)	A-Br (L)	C-B (L)
Present	−2.8	9.9	−1.3	14.2	−1.0	25.4	5.1	30.3
Martin (1968)	−1.9	25.8	13.2	20.4				

Note: H and L designate high and low meaningfulness respectively.

A comparison of the measures of second-list learning for the two types of list (trigram-adjective and adjective-trigram) shows that the scores of corresponding conditions were in general quite similar. Thus, variations in stimulus meaningfulness and in response meaningfulness had essentially parallel effects on performance. In overall analyses of the two transfer measures, neither the main effect of the type of list nor any of its interactions with other variables approached significance.

First-list item strength and transfer

The hypothesis of encoding variability implies an inverse relation, for individual stimuli in a list, between the strength of the first-list association and transfer performance. Presumably, ease of first-list learning, and hence strength at criterion, reflects the stability of stimulus encoding—stability at the same time conducive to associative interference in the transfer phase. The same negative correlation can, of course, be predicted on the assumption that the strength of the competing prior association determines the amount of associative interference for a particular pair. However, if encoding variability is the decisive factor, the correlation should be more pronounced when stimulus meaningfulness is low than when it is high. This further prediction follows from the premise that the intrinsic variability of encoding is greater in the former than in the latter case.

The relevant data are presented in Table 3. The first-list items of each subject in the A-Br groups were evenly divided into an 'easy' and a 'hard' category on the basis of the frequencies of correct anticipations during acquisition. The total numbers of correct responses during the 10 transfer trials were then determined for the corresponding second-list pairs (identical stimuli). The means of these scores are listed in the table. Clearly, there was no relation between the ease of first-list learning and transfer performance in the acquisition of the trigram-adjective lists at either level of stimulus meaningfulness. Thus, the prediction derived from the hypothesis of encoding variability was not confirmed. (In the case of the adjective-trigram lists, there was a slight correlation for low-meaningfulness pairs, but it was not reliable.)

Table 3. Mean numbers correct in transfer as a function of ease of first-list learning

Condition	Trigram-adjectives		Adjective-trigrams	
	Hard	Easy	Hard	Easy
A-Br (H)	18.7	18.8	18.7	18.9
A-Br (L)	16.7	17.7	13.4	17.3

Note: H and L designate high and low meaningfulness respectively.

Interlist intrusions

The greater the stability of encoding, the higher should be the probability that the first-list response to a given stimulus will continue to be elicited during second-list learning. Hence, the frequency of interlist intrusions in transfer should increase as a function of stimulus meaningfulness. A direct test of this deduction is difficult, since intralist errors and interlist intrusions cannot be differentiated under the A-Br paradigm. However, an estimate of the relative frequency of associative intrusions is provided by the ratio of previously correct responses to the total number of intralist errors. This ratio, based on frequencies pooled over subjects, was 75/188 = .39 when stimulus meaningfulness was high, and $58/128 = .45$ when it was low. Thus, the small difference that was found was in the direction opposite to the predicted one. The corresponding values were $41/128 = .32$ and $37/98 = .38$ for the high- and low-meaningfulness adjective-trigram lists respectively. The persistence during transfer of first-list responses to specific stimuli appears to be independent of both stimulus and response meaningfulness.

First-list recall

The mean numbers of first-list responses recalled after the end of interpolated learning are shown in Table 1. For the trigram-adjective lists, retention was lower under the A-Br than under the C-B paradigm [$F(1, 60) = 18.47, p < .01$]. This difference was greater for the high- than the low-meaningfulness lists, but the interaction of paradigm and meaningfulness was not significant ($F = 2.22$). The direction of the difference agrees with the trend in first-list relearning observed by Martin. However, the pattern of recall scores was the same for the adjective-trigram lists, yielding Fs of 4.05 ($p < .05$) for the difference between paradigms and of less than 1.00 for the interaction. In an overall analysis of the recall scores, F was less than 1.00 for all interactions involving type of list.

DISCUSSION

The purpose of this investigation was to reexamine the interaction of paradigm and stimulus meaningfulness that Martin interpreted as evidence for the operation of the principle of encoding variability in transfer. Both response and stimulus meaningfulness were manipulated in order to assess the merits of an alternative explanation in terms of differences in first-list strength related to meaningfulness. The original objective could not be accomplished in its entirety because Martin's results were not substantiated. The basic measures of transfer showed trends in the expected direction, but they fell far short of significance; and percentage measures like those used by Martin produced unequivocally negative results, that is, a complete absence of any interaction of paradigm and stimulus meaningfulness. The manipulations of response meaningfulness and of stimulus meaningfulness yielded essentially equivalent trends.

The differences in material and procedure may be responsible for the apparent discrepancy between the two studies. The important fact is, however, that under conventional conditions of practice (those used in the large majority of transfer studies), the critical interaction failed to materialize to any significant degree. It may not be amiss to suggest here that Martin's original evidence was sufficiently marginal to make the present negative results not too surprising. In his study the extent of the interaction in terms of the untransformed measures of acquisition was small, and there is no indication that it was significant. The differences seemed more impressive when the percentages of net transfer were considered, although (as already noted) the significance of the critical interaction was not reported and percentages taken to a widely varying base are apt to be unstable and should be viewed with caution. The purpose of these comments is not to press the statistical criticisms of Martin's analysis but rather to emphasize the point that the data in hand appear to have been marginal. Consequently, the present findings leave the original conclusion about the effects of stimulus meaningfulness on transfer with minimal empirical support. In view of the negative results of the studies mentioned in the introduction to this paper, it is fair to conclude that there is little evidence at the present time for encoding variability as a significant factor in transfer.

It was suggested that differences in first-list strength as a function of either stimulus or response meaningfulness might influence the amount of associative interference under the A-Br relative to the C-B paradigm. The measures of transfer gave no support to this speculation. It should be noted in this connection that Merikle's results (with paired trigrams and under different conditions of acquisition and testing; 1968) were not duplicated either. The exact conditions under which variations in stimulus meaningfulness or in response meaningfulness influence the amount of associative interference remain to be identified.

Notes

This research was supported by Grant MH 12006 from the National Institute of Mental Health. Received for publication December 11, 1970.

1. Except for one item with a value of 99%.

References

Archer, E. J. 1960. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs* 74(10, whole no. 497).

Goggin, J. and Martin, E. 1970. Forced stimulus encoding and retroactive interference. *Journal of Experimental Psychology* 84:131–136.

Goulet, L. R. 1965. Interlist response meaningfulness and transfer effects under the A-B, A-C paradigm. *Journal of Experimental Psychology* 70: 264–269.

Houston, J. P. 1967. Stimulus selection as influenced by degrees of learning, attention, prior associations, and experience with stimulus components. *Journal of Experimental Psychology* 73:509–516.

Jung, J. 1963. Effects of response meaningfulness (*m*) on transfer of training under two different paradigms. *Journal of Experimental Psychology* 65: 377–384.

Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review* 75:421–441.

Merikle, P. M. 1968. Paired-associate transfer as a function of stimulus and response meaningfulness. *Psychological Reports* 22:131–138.

Postman, L. 1962. Transfer of training as a function of experimental paradigm and degree of first-list learning. *Journal of Verbal Learning and Verbal Behavior* 1:109–118.

Weaver, G. E., McCann, R. L., and Wehr, R. J. 1970. Stimulus meaningfulness, transfer, and retroactive inhibition in the A-B, A-C paradigm. *Journal of Experimental Psychology* 85:255–257.

Williams, R. F., and Underwood, B. J. 1970. Encoding variability: Tests of the Martin hypothesis. *Journal of Experimental Psychology* 86:317–324.

continued from page 596

Goldstein, Arnold P. *Psychotherapeutic Attraction*. Elmsford, N.Y.: Pergamon Press, 1971. Pp. 250. $9.50.

Greenbaum, Sidney. *Verb-Intensifier Collocations in English*. The Hague: Mouton, 1970. Pp. 96.

Grinspoon, Lester. *Marihuana Reconsidered*. Cambridge, Mass.: Harvard University Press, 1971. Pp. xi, 443. $9.95; paperback, $2.45.

Haas, W. *Phonographic Translation*. Manchester, England: Manchester University Press, 1971. Pp. 94. $5.00.

Habenstein, Robert W., ed. *Pathways to Data: Field Methods for Studying Ongoing Social Organizations*. Chicago: Aldine, 1970. Pp. xi, 276. $7.95.

Heiby, Walter A. *Live Your Life: Maxims on the Fine Art of Living*. Chicago: Living Books, 1971, Pp. viii, 177. Paperback, $.95.

Heim, Alice. *Intelligence and Personality: Their Assessment and Relationship*. Baltimore: Penguin Books, 1970. Pp. 206. Paperback, $1.25.

Heine, Patrick Johns. *Personality in Social Theory*. Chicago: Aldine, 1971. Pp. x, 197. $7.50.

Hernandez, David E. *Writing Behavioral Objectives: A Programmed Exercise for Beginners*. New York: Barnes & Noble, 1971. Pp. 45. Paperback, $1.50.

Herron, William G.; Green, Mary; Guild, Maurice; Smith, Anne; and Kantor, Robert E. *Contemporary School Psychology*. Scranton: International Textbook, 1970. Pp. 270. $6.95.

Hill, John P., and Shelton, Jev, eds. *Readings in Adolescent Development and Behavior*. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xii, 446. Paperback, $7.50.

Hirschi, Travis. *Causes of Delinquency*. Berkeley and Los Angeles: University of California Press, 1971. Pp. 309. Paperback, $3.25.

Howard, Alan. *Learning to be Rotuman: Enculturation in the South Pacific*. New York: Teachers College Press, 1970. Pp. xii, 184. $7.25; paperback, $4.95.

Howells, John G. *Theory and Practice of Family Psychiatry*. New York: Brunner/Mazel, 1971. Pp. xv, 953. $25.00.

Jeffer, H. Paul, and Levitan, Dic. *See Parris and Die: Brutality in the U. S. Marines*. New York: Hawthorn Books, 1971. Pp. xiv, 242. $6.95.

Joyce, C. R. B., ed. *Psychopharmacology: Dimensions and Perspectives*. New York: Barnes & Noble, 1971. Pp. ix, 430. Paperback, $5.00.

Katona, George; Strumpel, Burkhard; and Zahn, Ernest. *Aspirations and Affluence*. New York: McGraw-Hill, 1971. Pp. vii, 239. $12.95.

Katz, Aaron Hillel. *Mission of Man*. New York: Philosophical Library, 1970. Pp. 104. $5.00.

Kennedy, Wallace A. *Child Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xxxiv, 604. $10.50.

Kimmel, H. D. *Experimental Principles and Design in Psychology*. New York: Ronald Press, 1970. Pp. vi, 280. $8.50.

King, Richard A., ed. *Readings for an Introduction to Psychology*. 3d ed. New York: McGraw-Hill, 1971. Pp. viii, 584. Paperback, $4.95.

continued on page 500

American Journal of Psychology
1971, Vol. 84, No. 4

'CODING' IN STUDIES OF ACOUSTIC AND SEMANTIC INTERFERENCE AND CONFUSION

Michael M. Gruneberg and Robert N. Sykes
University College, Swansea, Wales

This paper considers what is meant by the term 'coding' in those experiments on acoustic and semantic interference and confusion that claim to demonstrate different coding systems in short- and long-term memory. It is questioned whether any inference as to the nonexistence of a coding system can ever be made on the basis of interference and confusion studies.

One of the main lines of evidence claimed to support a distinction between short- and long-term memory systems comes from the studies of immediate memory showing that acoustically related sequences are less well recalled than are semantically related sequences or nonrelated sequences (Baddeley, 1966a). From this evidence it is concluded that semantic coding does not occur in short-term memory. It was also concluded by Baddeley (1966b), on the basis of experiments on the effects of acoustic and semantic similarity of items on long-term retention, that acoustic coding is confined to short-term memory. Two questions arise from this type of experiment. What is meant by coding? And what conclusions about coding can be drawn from experiments on interference and confusion effects?

Specifically, when one deals with the meaning of the term 'coding' in the context of acoustic-confusion experiments, two possibilities are open. By coding might be meant (*a*) that each item is represented along an acoustic dimension or (*b*) that the *relationships* between items are stored along an acoustic dimension. Now, as Gruneberg, Colwill, Winfrow, and Woods (1970) have pointed out, the fact that one can reproduce on demand words which are acoustically related to stimulus words indicates long-term acoustic storage. This is because no matter what strategy one employs in generating acoustically similar words, there must be a point at which the internal representation of the phonemic sequence generated is compared with stored material to confirm that the sequence *is* a word. At the very minimum, therefore, it is necessary to postulate acoustic storage of items over the long term in order to account for this type of phenomenon.

Again, while it is *logically* possible that the representation of the acoustic relationship between items is not formed, or has atrophied over the long term, the fact that one can produce, on request, acoustically similar words indicates at least the possibility of acoustic associations in memory over the long term. Those who wish to establish a dichotomous theory of memory on the basis of differential coding must deny a role to acoustic relationships over the long term. It is not enough to point out that there are other possibilities than acoustic relationships to account for the evidence; it must be positively shown that the assumption of acoustic relationships is not justified.

The second question concerns the inferences that can be drawn from experiments on confusion and interference effects. If one were to claim, as Baddeley (1966a) reasonably might, that by coding is meant relationships between items rather than any coding of the items themselves (in Baddeley's experiment the actual items were available for inspection, the subjects having only to recall the order in which the items were presented), then the assumption must be that confusion effects indicate underlying coding of relationships, whereas the failure to find confusion effects allows one to presume that the coding of relationships does not exist.

Yet this last presumption is clearly not logically defensible. Lack of confusion effects for, say, semantically related material might merely indicate that material related in such a manner is more resistant to the effects of certain factors which cause forgetting. It cannot be concluded that underlying semantic relationships do not exist.[1] Hence it cannot be concluded on the basis of experiments involving confusion or interference effects that the absence of such effects indicates the absence of certain types of coding systems.

It might be added that this second argument is true also of evidence claiming to show a precategorical acoustic store (Morton, 1970), for interference-effects evidence of an acoustic store does not preclude the possibility of the existence of other stores undetectable by this method. As material being listened to can almost immediately be identified in terms of some semantic categories (e.g., language; 'it is English,' or 'German,' or whatever), it is reasonable to doubt the adequacy of the conclusion that there is a precategorical acoustic store which lasts more than a fraction of a second. To do so on the basis of experiments on interference effects is analogous to arguing that because Woolworth's can be shown to sell sweets it is a sweetshop.

Although the claim that short- and long-term memory involve different coding systems hardly seems justified by evidence of interference and confusion effects, it is still theoretically possible that acoustic interference or confusion effects are confined to the short term and semantic interference or confusion effects to the long term. The empirical evidence is, however, against such a hypothesis. Underwood and Freund (1968), Gruneberg and Sykes (1969), and Gruneberg et al. (1970) have shown acoustic-confusion effects over the long term using a recognition test. Wickens et al. (1970) and Bruce and Crowley (1970) have shown acoustic interference effects in long-term memory using recall techniques, and Baddeley himself (Baddeley and Warrington, 1970) allows that an earlier experiment of Brown and McNeil (1966) indicates acoustic coding in long-term memory. Brown and McNeil found that when subjects guessed a word for which a meaning was supplied, an acoustically similar but incorrect word was often given.

As to semantic confusion in short-term memory, evidence of this is present in the work of Underwood and Freund (1968), using a recognition technique, and in that of Loess (1968) and Corman and Wickens (1968), using recall techniques. Some effects of semantic confusion in short-term memory are also to be found in the original work of Baddeley (1966a).

The empirical evidence, then, also is against establishing a dichotomous theory of memory on the basis of certain 'characteristic' types of interference or confusion effects in one memory system rather than another. It seems reasonable to suppose that incoming items are coded both as individual items and relationally, in terms of both sensory and semantic characteristics, and that such coding is persistent over the long term.

Notes

Received for publication October 5, 1970.

1. However, studies that show semantic coding in the short term using release from proactive inhibition (Wickens, 1970) can probably claim only to show *item* coding, not relational coding.

References

Baddeley, A. D. 1966a. Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology* 18:362–365.

Baddeley, A. D. 1966b. The influence of acoustic and semantic similarity on long-term memory for words. *Quarterly Journal of Experimental Psychology* 18:302–309.

Baddeley, A. D., and Warrington, E. K. 1970. Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior* 9:176–189.

Brown, R., and McNeil, D. 1966. The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior* 5:325–337.

Bruce, D., and Crowley, J. J. 1970. Acoustic similarity effects on retrieval from secondary memory. *Journal of Verbal Learning and Verbal Behavior* 9:190–196.

Corman, C. D., and Wickens, D. D. 1968. Retroactive inhibition in short-term memory. *Journal of Verbal Learning and Verbal Behavior* 7:16–19.

Gruneberg, M. M., and Sykes, R. N. 1969. Acoustic confusion in long-term memory. *Acta Psychologica* 31:293–296.

Gruneberg, M. M., Colwill, S. J., Winfrow, P., and Woods, R. W. 1970. Acoustic confusion in long-term memory: An extension of previous findings. *Acta Psychologica* 32:394–398.

Loess, H. 1968. Short-term memory and item similarity. *Journal of Verbal Learning and Verbal Behavior* 7:87–92.

Morton, J. 1970. The role of memory in the attention process. *Acta Psychologica* 33:271–279.

Underwood, B. J., and Freund, J. S. 1968. Errors in recognition learning and retention. *Journal of Experimental Psychology* 78:55–63.

Wickens, D. D. 1970. Encoding categories of words: An empirical approach to meaning. *Psychological Review* 77:1–15.

Wickens, D. D., Ory, N. E., and Graf, S. A. 1970. Encoding by taxonomic and acoustic categories in long-term memory. *Journal of Experimental Psychology* 84:462–469.

American Journal of Psychology
1971, Vol. 84, No. 4

THE TWO ILLUSIONS OF MÜLLER-LYER: CONFUSION THEORY REEXAMINED

Robert Sekuler and Albert Erlebacher
Northwestern University

The Müller-Lyer illusion was studied in both the tails-in and tails-out forms, with both the distance between the tips of the tails and the size of the angle formed by the tails and the shaft factorially varied. The results force two conclusions: that there are two functionally different Müller-Lyer illusions, one with tails in and one with tails out, and that the confusion theory offered earlier does not explain both.

'Confusion' theories (Woodworth, 1938, p. 645) have been prominently used to explain the Müller-Lyer and other geometrical illusions. Erlebacher and Sekuler (1969) addressed the question of what, in fact, might be 'confused' when subjects judge the length of lines with appended 'tails,' or obliques. We concluded that subjects judge not only the actual line length but incorporate into their judgment the distance between the ends of the obliques. The perceived length then is a compromise between these two quantities. This conclusion resulted from a series of experiments on that version of the Müller-Lyer figure with the obliques directed inward. The independent variables were the length of obliques and the angle these obliques made with the shaft.

Briefly, those experiments showed (*a*) that with the size of the angle between obliques and shaft varied, and with the lengths of those obliques held constant, then the smaller the angle, the smaller the shaft appeared and the greater the illusion; (*b*) that with the size of the angle varied, and with the lengths of the obliques varied to keep a constant distance between their ends, then the magnitude of the illusion remained constant; and (*c*) that with the size of the angle and the distance between the ends of the obliques factorially varied, then there was a large effect of distance (the smaller distances causing the greater illusions), a very small and statistically nonsignificant effect of angle size, and a small but significant interaction of distance and angle size. The interaction suggested revising the basic notion that the illusion is caused by the distance between the ends of the obliques—revising it to reflect the fact that when the ends of the obliques are near (perpendicular distance) the to-be-judged line, they have a greater effect than when they are farther away.

The present study was designed to test the confusion hypothesis on that portion of the Müller-Lyer figure with the tails directed outward. Again, the distance between the ends of the obliques (intertip distance) and the angle the obliques made with the shaft of the figure were varied in a factorial design. Our original observations (Erlebacher and Sekuler, 1969) on the tails-in Müller-Lyer figure were repeated in the present experiment to facilitate comparison between the data for both forms of the illusion.

To further broaden the base of generalizability of the previous results, several other changes were made. In our earlier experiment, the method of reproduction was used to obtain points of subjective equality (PSEs). In the present experiment, the up-and-down method of Dixon and Mood (1948) was employed. Moreover, in the earlier experiment the figures were rather large—averaging about 4.5 deg visual angle. In the present experiment the figures subtended approximately 1.5 deg visual angle.

METHOD

—Design—PSEs were obtained for 42 figures. Of these, 33 were experimental figures with a 43-mm shaft length. Nine 'filler' figures whose shaft length was 46 mm were used so that the subject would have nonillusory length differences available, increasing the apparent validity of his task. The experimental figures were of three types. Sixteen figures were of the tails-in variety and came from a factorial combination of four angle sizes (10, 30, 50, and 70 deg) and four intertip distances (12.90 mm, 19.35 mm, 25.80 mm, and 32.25 mm). Sixteen more figures were of the tails-out type and came from the factorial combination of four angle sizes (170, 150, 130, and 110 deg) and four intertip distances (73.10 mm, 66.65 mm, 60.20 mm, and 53.75 mm). One figure was a plain line with no tails. It was introduced as a reference stimulus to evaluate any constant error not caused by the illusion-producing obliques. Of the filler figures, four were of the tails-in variety, four of the tails-out type, and one a simple straight line.

—Subjects—Twenty-three male and female semi-volunteers, students in introductory psychology at Northwestern University, served in all conditions of the experiment. They were individually tested.

—Stimuli—The 42 standard stimuli were vertically oriented black-on-white figures rear-projected onto a film-coated glass screen by a film-strip projector. The illuminated area measured 21.5 cm vertically by 16 cm horizontally and the figures were centered in this space. The shaft length was 43 mm for the experimental figures and 46 mm for the filler figures. The length of the obliques was adjusted for each figure in order to obtain the specified angle of obliques with shaft and the specified intertip distance. The background of the figures had a luminance of 136 cd/m² and the contrast of the stimuli was approximately .90. The subject sat 1.75 m from the screen, which was about at eye level. These geometrical relationships yielded figures whose straight-line segments subtended 1 deg 24 min and 1 deg 30 min for the experimental and filler stimuli respectively. The 42 standard stimuli were in a strict random order, which repeated itself in a continuous loop after each presentation of all 42 figures. Two such random orders were used, 8 subjects being run with one and 15 with the other.

The variable stimulus required by the psychophysical method employed was a vertically oriented line without tails, presented in the center of a 10-cm. computer-controlled cathode-ray tube. The line was lighted while the background was dark. The luminance of the line was 17 cd/m². The center of the cathode-ray tube was 48 cm to the left of the center of the screen on which the standard stimuli were projected, and it was displaced 17.5 cm upward to avoid all-too-reliable vernier comparison.

—Procedure—The psychophysical method used was the up-and-down method (Dixon and Mood, 1948). For any one stimulus the method was as follows. On the first presentation of a standard stimulus, the variable stimulus was 41.5 mm long for all stimuli. The subject then indicated whether the standard appeared longer or shorter than the variable. If the standard was judged as longer, then on the next presentation of that standard, the variable was made 1.5 mm longer, or 43 mm in length. If the standard was judged as shorter than the variable, then on the next trial that this standard was presented, the variable stimulus was made 1.5 mm shorter, or 40 mm long. On each succeeding presentation of that standard stimulus, the variable stimulus was made 1.5 mm longer or shorter, depending on whether on its last presentation that standard was judged as longer or shorter than the variable. There were 11 trials for each of the 42 standard stimuli, making a total of 462 judgments. The subject indicated his judgment by flicking a toggle switch to the left or right.

Giving all 11 trials for a given standard stimulus all at once would have produced a spurious correlation among the judgments. Indeed, this could occur whenever the subject could remember what response he had made on the last presentation of a particular stimulus. To avoid such a bias, the second trial on a given stimulus came after the subject had made a judgment for each of the other 41 stimuli. In other words, 41 judgments intervened between successive judgments involving a given standard stimulus. It would be very difficult, if not impossible, to identify such a stimulus from among 42 different ones (Miller, 1957), and even if such identification were possible the subject would be hard put to remember what response he had made to it 42 trials earlier. The assumption can be made, then, that each trial yielded an independent judgment.

Timing of stimulus presentation was as follows: first a standard stimulus appeared on the screen, and .15 sec later the variable stimulus appeared on the cathode-ray tube. Both stimuli disappeared simultaneously, 3.15 sec after the onset of the standard stimulus (and 3.00 sec after the onset of the variable). The subject could make his response only after the offset of the stimuli. Though he had an unlimited time to make his response, the average response time was 1 sec. A new standard stimulus appeared 1.75 sec after the response was made. Since the screen of the cathode-ray tube was coated with a short-persistence (P-31) phosphor, the standard stimulus became invisible within several milliseconds of the time at which the computer ceased to refresh the display.

RESULTS

Calculation of PSEs

A PSE was obtained for each subject for each standard stimulus by the usual procedure for the up-and-down method. For the 11 trials on which a given standard stimulus was presented, it was determined which response—'greater' or 'lesser' than the variable—was the less frequent. The mean length of the variable stimulus on those trials when that response was made was then calculated. If that less frequent response was 'greater,' then .75 mm (half the 1.5-mm step size) was added to the calculated mean. If the less frequent response was 'lesser,' then .75 mm was subtracted from the calculated mean.

Tails-in illusion

The mean PSEs for the tails-in figures as a function of intertip distance are shown in the bottom half of Figure 1. Also plotted in Figure 1 is the 41.39-mm PSE for the line without tails. Figure 1 shows that as the intertip distance for the tails-in figures increased, so did the PSE. The magnitude of the illusion decreased. That same relationship can be seen in Figure 2, where angle size is shown on the abscissa and intertip distance is the parameter. Moreover, it can be seen in Figure 2 that, in general, as angle size increased, so did the mean PSE. The 70-deg figures yielded a very small illusion; the other angle conditions, differing slightly from each other, yielded a larger illusion. An analysis of variance showed that both intertip distance and angle size were significant sources of variation [$F(3, 66) = 17.25, p < .01$, and $F(3, 66) = 23.85, p < .01$, for distance and angle size respectively]. The interaction between these two variables was not statistically significant [$F(9, 198) = 1.79$]. The mean PSEs for the four angle sizes were 38.54 mm, 38.84 mm, 39.25 mm, and 40.64 mm for the 10-, 30-, 50-, and 70-deg angles respectively.

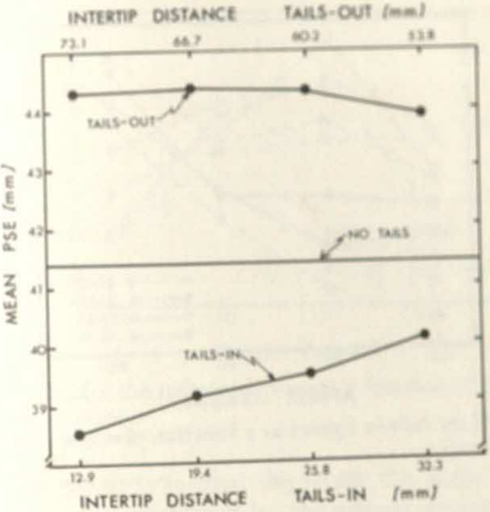


Fig. 1. Mean PSE for both tails-in and tails-out figures as a function of intertip distance; the PSE for the tailless line is also given

Tails-out illusion

The top portion of Figure 1 shows the mean PSEs for the tails-out figures as a function of intertip distance. The abscissa values are indicated at the top of the figure and are shown in decreasing magnitude. It can be seen that intertip distance had virtually no effect on the PSE.

The reason for inverting the axis in the top part of Figure 1 was to facilitate comparison with experiments in which the entire Müller-Lyer figure is used—experiments in which there is a line with tails inward as well as one with tails outward. In those cases, the two portions of the figure are 'symmetrical' (in that the tail lengths are equivalent and the angles for the tails-out portion and the tails-in portion are complementary) and a small intertip distance for the tails-in figure corresponds to a large intertip distance for a tails-out figure. Accordingly, Figure 1 places data points for corresponding tails-in and tails-out figures in the same column. If one assumed additivity of the two portions of the illusion, one might predict the amount of illusion by the vertical distance between corresponding points in the two parts of Figure 1. In this fashion, one might say that the largest illusion would result from those figures with the longest tails. But this relationship is due largely to the intertip

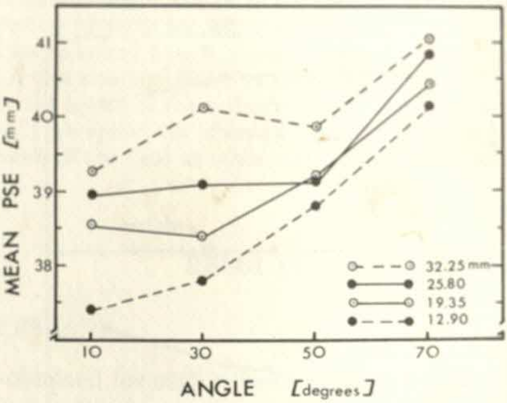


Fig. 2. Mean PSE for tails-in figures as a function of angle size; intertip distance is the parameter

distance of the tails-in figure, since the tails-out PSEs varied but little with intertip distance.

Finally, Figure 3 shows the mean PSEs as a function of angle size. Intertip distance is the parameter. Angle size is given in decreasing order on the abscissa for comparability with respect to the full two-part Müller-Lyer figure, as explained above. Figure 3 indicates that angle size had a nonmonotonic effect on the PSE. Averaged over intertip distances, the mean PSEs were 43.41 mm, 44.46 mm, 45.02 mm, and 43.94 mm for the 170-, 150-, 130-, and 110-deg angles respectively. Analysis of variance indicated that intertip distance was not a significant source of variation [$F(3, 66) = 1.13$]. Angle size, however, did have a significant effect [$F(3, 66) = 5.27, p < .01$ or $p < .05$].[1] The interaction was perhaps significant [$F(9, 198) = 2.33, p < .05$ or $p < .15$].

DISCUSSION

The data from the present experiment are generally consistent with the earlier data (Erlebacher and Sekuler, 1969) on the *tails-in* form of the Müller-Lyer figure. The results with the tails-in figures of the present experiment confirm the finding, for example, that the intertip distance is an effective variable determining the size of the constant error.

No significant effect of angle size per se was found in the earlier work, but it was predicted that such an effect would be manifest given sufficiently sensitive measurements and a sufficient range of stimulus

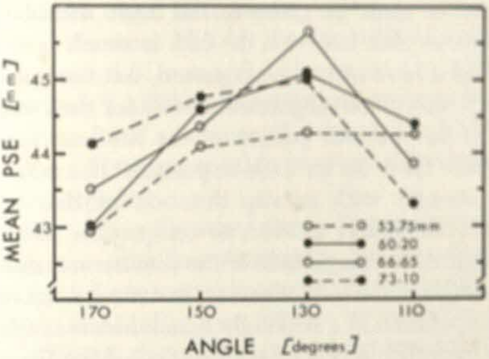


Fig. 3. Mean PSE for the tails-out figures as a function of angle size; intertip distance is the parameter

variation. That prediction—that the larger the angle, the smaller the illusion—was in fact confirmed by the present results. The reasoning behind the prediction was that basically the illusion is caused by the subject's incorporating the intertip distance into his judgment of the length of the shaft of the figure. However—that reasoning continues—with intertip distance fixed, the variation of angle size produces a concomitant variation in the perpendicular distance between the ends of the tails and the line to be judged. And as that perpendicular distance increases, the illusion-producing portion of the figure becomes more distant from the line itself and the magnitude of the illusion should then be diminished.

The only fact from the present experiment not in accord with that set of arguments is that the present experiment yielded no significant interaction between intertip distance and angle size. This interaction was found in the earlier study and attributed to the variation in the distance of the end of the tails from the shaft of the figure. There now appears to be an independent effect of angle size as well as the effect of intertip distance.

Although this discrepancy between the two experiments is puzzling, there are far too many differences, in the parameters of the stimuli, that cannot be discounted as possible sources of the discrepancy. For example, the present experiment used stimuli subtending far smaller visual angles than the previous experiment. It was hoped that such reduced size would eliminate or minimize the effects of attention and of multiple partial looks at the stimuli (Pollack, 1970). If this was a successful attempt, the data of this experiment are in some sense 'better.' In general, much greater notice must be given to the basic stimulus parameters in studies of illusions than has been the case in much of the past literature. Such variables have often been ignored, but the work of Wickelgren (1965) and Fisher (1969) argues effectively for their importance.

The results on the *tails-out* portion of the Müller-Lyer figure were very different from those on the tails-in portion. The magnitude of the illusion varied strongly with intertip distance for the tails-in figures, whereas the tails-out figures showed no comparable effect. Moreover, for the tails-in figures the magnitude of the illusion was monotonic with variation in the angle of the tails, whereas the tails-out figures responded to the same manipulation in a strikingly nonmonotonic fashion. Though it is difficult to formulate a precise explanation of the illusion resulting from the tails-out figures, the phenomena observed here must be accounted for if one is to understand the operation of what is referred to as the Müller-Lyer illusion.

One fact emerges strongly from the present experiment: what is usually called *the* Müller-Lyer illusion is actually two different illusions. Both this experiment and others indicate that the illusions resulting from the tails-in and the tails-out figures respond in quite different ways to certain stimulus manipulations. For example, Piaget (1969) found that with tachistoscopic presentation, variation in the exposure time of the figures affected the magnitude of the illusion from the two types of figures differentially. With the tails-out version there was a monotonic relationship—the shortest exposure time producing the greatest illusion. The relationship was not systematic for tails-in figures. Furthermore, Piaget found that with varying intertip distance (and covarying angle size, to achieve a constant perpendicular distance of the ends of the tails to the shaft of the figure), there was a monotonic relationship for tails-in figures like the one found in the present experiment—the greatest illusion resulting from the smallest intertip distance. The relationship with tails-out figures, however was nonmonotonic.

Some observations reported by Runyon and Cooper (1969) also support the dual-illusion hypothesis. They found that removing the shaft (leaving only the tails) from the tails-out figures decreased the magnitude of the illusion. Performing the same operation on the tails-in figures had the opposite effect.

Finally, there is the well-known fact (Binet, 1895; Heymans, 1896) that the absolute size of the illusion produced by the tails-in portion of the Müller-Lyer figure is less than that produced by the tails-out portion. In the present experiment the mean PSE for the tails-in figures was

39.32 mm, and the mean PSE for the tails-out figures was 44.20 mm. Measuring from the mean PSE for the line without tails (41.39 mm), the mean illusions were 2.07 mm and 2.81 mm for the tails-in and tails-out figures respectively.[2] All these observations support the idea that after nearly three-quarters of a century of work on the Müller-Lyer illusion, we find ourselves in the unenviable position of having not just one but two illusions to explain.

The existence of two illusions, responding in quite different ways to variation in certain stimulus parameters, has two major consequences, one theoretical and one somewhat more practical. Any theory that purports to explain the Müller-Lyer illusion by means of a single factor is necessarily incorrect, since there are at least two functionally separate, and often incompatible, sets of data to be explained, one for tails-in and one for tails-out. An example is the 'depth-cue scaling' theory most recently championed by Gregory (1968) and most strongly supported by evidence from experiments in which the perceived depth of a luminous Müller-Lyer figure is compared with the perceived size of the figure; the experimental parameter is the angle of the tails. As one might expect from the dual-illusion hypothesis, this scaling theory is differential in its ability to handle the data from the tails-in and tails-out figures: from Gregory's published curves (1968, p. 74), the correlation between perceived size and perceived depth seems far better for the tails-in figures than for the tails-out ones. Similar theoretical obstacles necessarily hold for theories that attribute the Müller-Lyer illusions to univariate neural processes such as lateral inhibition (Bekesy, 1967) or the operation of orientation analyzers (Burns, 1968, pp. 104–105). The nonmonotonic effect of angle size and the absence of a strong effect of intertip distance with the tails-out figures is inconsistent with these theoretical approaches.

In practical terms, the existence of two illusions means that experimenters should avoid the unnecessary complications introduced by any single figure that combines both illusions, at least until the details on each illusion are independently sorted out and the laws governing their interaction are understood. Just such a figure, often called the Brentano version of the Müller-Lyer illusion, has been commonly used in the past and is found, in one or another of its commercial forms, in many experimental psychology laboratories. Our data, and the existence of two Müller-Lyer illusions, would make the Brentano version of the illusion seem far too complicated for the present pitiful state of the art of understanding illusions.

Notes

Computer time was made possible through Grant EY 00321 from the National Institutes of Health. The first author's complete address is: Department of Psychology, Northwestern University, Evanston, Illinois 60201. The paper was received for publication January 13, 1971.

1. Because of repeated measurements on the same subject, the significance level is not clear. The lower level given is the appropriate one if there is no correlation of error components; the higher value given is the appropriate one for perfect error correlation (see Greenhouse and Geisser, 1959). The correct value is somewhere between these.
2. Although most studies agree fairly well on the relative magnitudes of tails-in and tails-out illusions, the present data suggest that those relative magnitudes should not be assumed constant and subject only to the vagaries of experimental error. Rather, the relative magnitude depends on the particular values of angle and intertip distance used in any given experiment (see Figure 1).

References

Binet, A. 1895. La mesure des illusions visuelles chez les enfants. *Revue philosophique de la France et de l'étranger* 40:11–25.

Bekesy, G. von. 1967. *Sensory inhibition*. Princeton, N.J.: Princeton University Press.

Burns, B. D. 1968. *The uncertain nervous system*. London: Edward Arnold.

Dixon, W. J., and Mood, A. M. 1948. A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association* 43:109–126.

Erlebacher, A., and Sekuler R. 1969. Explanation of the Müller-Lyer illusion: Confusion theory examined. *Journal of Experimental Psychology* 80: 462–467.

Fisher, G. H. 1969. Towards a new explanation for the geometrical illusions. *British Journal of Psychology* 60:179–185.

Greenhouse, S. W., and Geisser, S. 1959. On methods in the analysis of profile data. *Psychometrika* 24:95–112.

Gregory, R. L. 1968. Perceptual illusions and brain models. *Proceedings of the Royal Society* (series B) 171:279–296.

Heymans, G. 1896. Quantitative Untersuchungen über das "optische Paradoxon." *Zeitschrift für Psychologie und der Physiologie der Sinnesorgane* 9:221–255.

Miller, G. A. 1956. The magical number seven, plus or minus two. *Psychological Review* 63:81–97.

Piaget, J. 1969. *The mechanisms of perception*. New York: Basic Books.

Pollack, R. H. 1970. Mueller-Lyer illusion: Effect of age, lightness, contrast, and hue. *Science* 170:93–95.

Runyon, R. P., and Cooper, M. R. 1970. Enhancement of Sander illusion in minimal form. *Perception and Psychophysics* 8:110–111.

Wickelgren, B. G. 1965. Brightness contrast and length perception in the Müller-Lyer illusion. *Vision Research* 5:141–150.

Woodworth, R. S. 1938. *Experimental psychology*. New York: Holt.

American Journal of Psychology
1971, Vol. 84, No. 4

VISUAL AND HAPTIC ANGLE PERCEPTION IN THE MATCHING TASK

Stuart Appelle
State University of New York at Brockport

Forty subjects reproduced angles of 30–150 deg from a series of comparison angles presented in ascending and descending orders. Magnitude of error was significantly lower for the visual-visual group than for the haptic-haptic, visual-haptic, and haptic-visual groups. The results suggest that haptic percepts of angle are smaller in magnitude than those for vision and that performance differences between modalities reflect both the nature of form discrimination and the nature of the matching task itself.

The present study had two purposes. First, it examined discrimination for a particular feature of form (namely, angularity) rather than form discrimination per se. Second, it examined the nature of intra- and crossmodal matching experiments. The argument presented is that performance differences between modalities reflect both the nature of form discrimination and the nature of the matching task itself.

The concern with particular features of form was based on theoretical considerations suggesting differences in the use of specific features by eye and hand (Revesz, 1950) and on several experiments reporting differences between vision and touch in the perception of specific aspects of form (Appelle and Goodnow, 1969; Cashdan, 1968; Walk, 1965). The nature of the matching experiment provides additional justification for an emphasis on specific features, as Goodnow (1969) has demonstrated that vision and touch may sample different properties of the same stimulus and that stimulus properties may differ in salience to the inspecting modalities. Thus, the extent of stimulus overlap when one inspects an object by eye and hand may account for differences between intra- and crossmodal matches (Goodnow, 1970).

Concern with particular properties of form is just one essential part of interpreting from intra- and crossmodal data. Other characteristics intrinsic to the matching experiment must also be considered. Björkman (1967), for example, has shown that intramodal matching may typically be more accurate than crossmodal matching due to the higher correlation of subjective impressions when standard and comparison are inspected within the same perceptual system. This should suggest attention to the subjective values of both standard and comparison objects, but few investigators take that approach when interpreting their results.

The choice of angle as the particular property to be studied arose out of several considerations. First, it stemmed from experiments indicating that angularity is an important feature in judgments of form complexity (Stenson, 1966; Brumaghim, 1967; Brown and Andrews, 1968; Brown and Brumaghim, 1968; Brumaghim and Brown, 1968) and that performance is different for visual (Jastrow, 1893; Pratt, 1926; Logan, 1964), haptic (Calabresi, 1933), and crossmodal (Lederman and Taylor, 1969) judgments. Second, attempts to resolve confusion over past angle experiments (Fisher, 1969) do not seem satisfactory. Finally, several experiments suggest that judgments may be influenced by the direction in which the judged angle moves (Pratt, 1926; Calabresi, 1933; Lederman and Taylor, 1969). This last factor is interesting in its own right. But also, it has generally operated only on comparison objects, not standards. It can be used, therefore, to illustrate the importance of considering judgments of both objects when interpreting from matching experiments.

METHOD

—Subjects—The subjects were male and female volunteers, students at George Washington University who were unaware of the purpose or nature of the experiment.

—Apparatus—The apparatus consisted of two pairs of wooden arms, 16.5 cm long and .5 cm wide, mounted on a Masonite board 48 cm long and 48 cm wide. Each pair of arms was constructed with one stationary horizontal arm and one variable arm that could form any angle from 0 to 180 deg. These 'angles' were constructed to increase clockwise in arc and were arranged one directly above the other, in identical orientation and separated by a distance of 25 cm between their horizontal arms. The Masonite mount was placed flat on a tabletop directly in front of the subject. A scale calibrated in degrees formed a semicircle just beyond the angle, giving each setting exactly. Comparison judgments were made on the angle closer to the subject, while the standards were always presented with the higher angle.

—Procedure—The subjects were randomly assigned to one of four experimental groups, 10 to each group. In the *haptic-haptic* group, both standard and comparison angles were presented haptically to blindfolded subjects. Haptic exploration was restricted to the angle arms themselves rather than to the angle's arc or complement, but the kind of haptic activity was unrestricted. The only qualification was that the preferred (writing) hand be used. In the *visual-visual* group, the subjects visually inspected both the standard and comparison angles. In the *haptic-visual* group, blindfolded subjects haptically explored a standard angle, then made a visual comparison. In the *visual-haptic* group, the subjects visually explored a standard angle, then were blindfolded and made a haptic comparison.

Three acute angles (30, 45, 60 deg), three obtuse angles (120, 135, 150 deg), and a 90-deg angle were used as standards. The subjects were instructed to explore each angle, either visually or haptically, until it was firmly fixed in their memory. They were then presented with a comparison angle, either visually or haptically, and asked to judge whether the comparison was larger than, smaller than, or the same size as the standard. If the comparison angle was judged to be larger or smaller than the standard, the comparison angle was adjusted by the experimenter, under the direction of the subject, until a point of subjective equality (PSE) was reached. Adjustment was continuous rather than discrete, and both visual and haptic inspection of the comparison angle were continuous and concurrent with the adjusting movement. The standard stimulus was always hidden from view during visual inspection of a comparison angle, nor could it be felt during haptic inspection of the comparison. This ensured that in all experimental groups, standard and comparison stimuli were always presented successively. Two judgments were obtained for each of the seven angles, one with the comparison stimulus presented in an ascending pattern (starting point at a value less than the objective value of the standard) and the other with the comparison stimulus presented in a descending pattern (starting point greater than the objective value of the standard). Thus, 14 judgments were made by each subject in each group. Starting points varied from 2 to 20 deg from the objective value of the standard. Order of presentation was randomized for all 14 trials.

RESULTS

The mean settings in degrees (PSEs) for each condition and angle are presented in Figure 1. Judgments made with an ascending comparison angle were always lower than judgments made with the corresponding descending comparison angle. Analysis of variance of these data revealed a significant interaction of comparison series and treatment [$F(3, 36) = 3.36, p < .05$]. Tests for simple effects were therefore carried out.

Differences among intra- and crossmodal groups were examined at both types of comparison series. The four groups differed significantly, both with the ascending [$F(3, 72) = 3.27, p < .05$] and descending [$F(3, 72) = 7.03, p < .01$] series of comparisons. A Newman-Keuls post-test showed that haptic-visual judgments made on ascending comparisons were significantly smaller than visual-haptic and visual-visual judgments, and significantly smaller than all other groups when judgments were made on descending comparisons ($p < .05$).

The effect of comparison series (ascending versus descending trials) was determined for each of the four experimental (modal) conditions.

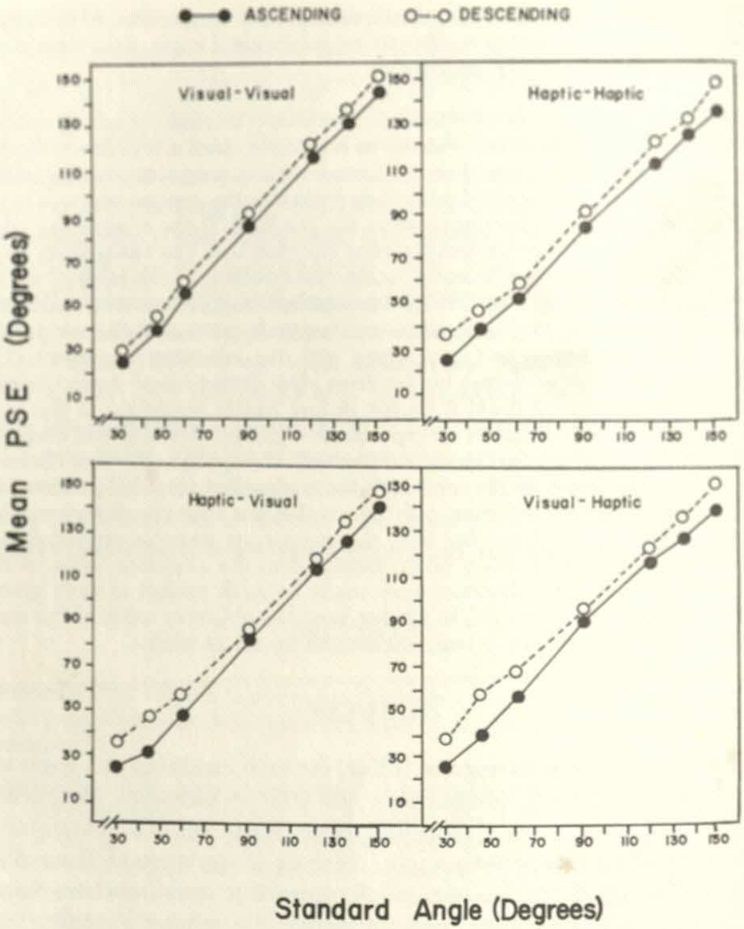


Fig. 1. PSEs for ascending and descending comparison series

Ascending and descending judgments were significantly different for each experimental group ($p < .01$). The corresponding values for $F(1, 36)$ were 53.9 (haptic-haptic), 16.82 (visual-visual), 30.82 (haptic-visual), and 69.73 (visual-haptic). Judgments on ascending comparisons were smaller than judgments on descending comparisons on 88.9% of all trials. In addition, they were almost always smaller than the true size of the standard. Descending judgments were almost always larger.

The analysis of variance also revealed a significant interaction between comparison series and angle $[F(6, 216) = 4.14, p < .01]$. Tests

Table 1. Mean absolute error (degrees) for comparison series ascending (A), descending (D), and A and D combined

	Standard angle (degrees)														
	30		45		60		90		120		135		150		Overall mean group error
Group	A	D	A	D	A	D	A	D	A	D	A	D	A	D	
Haptic-haptic	3.4	7.2	5.4	5.1	6.3	5.4	6.0	5.9	5.9	7.5	8.2	6.9	9.6	6.4	
Visual-visual	2.5	2.6	7.1	4.8	3.2	3.3	1.2	1.8	3.9	3.5	4.1	4.2	4.1	4.1	
Haptic-visual	5.8	7.0	10.8	6.8	11.1	7.4	7.0	6.3	6.5	5.2	12.2	7.5	11.2	9.8	
Visual-haptic	9.4	9.2	5.1	14.0	6.7	9.9	7.4	7.9	5.4	5.8	7.6	4.9	9.9	5.7	
A and D combined															
Haptic-haptic	5.3		5.3		5.9		6.0		6.7		7.6		8.0		6.37
Visual-visual	2.6		5.9		3.3		1.5		3.7		4.2		4.1		3.60
Haptic-visual	6.4		8.8		9.3		6.7		5.9		9.9		10.5		8.19
Visual-haptic	9.3		9.6		8.3		7.7		5.6		6.3		7.8		7.78

for simple effects were again carried out. Judgments were significantly smaller [$F(1, 216), p < .01$] on ascending than descending comparisons for all angles except 90-deg ones ($F = 3.59, p > .05$). Nevertheless, 75% of all subjects still set the ascending comparison for the 90-deg standard at a value lower than the descending one.

The data have been considered so far in terms of PSE. To give a better indication of accuracy, the results were also considered in terms of amount and direction of error. Table 1 shows the mean absolute error (ignoring direction of error). The data in the lower half of the table—the comparison series combined—suggest that the overall error was different among the four modal conditions. This was confirmed by an analysis of variance [$F(3, 36) = 10.67, p < .01$]. A Newman-Keuls posttest showed that the magnitude of error in the visual-visual condition was significantly lower than all other groups ($p < .01$). No other differences were significant, but it might be pointed out that error was greatest for the two crossmodal conditions.

Although the haptic-haptic group showed a consistent tendency for error to increase with increasing angle size (Table 1, comparison series combined), no consistent pattern emerged from the other conditions. The analysis of variance revealed significant differences in magnitude of error among angles [$F(6, 216) = 2.36, p < .05$], but a Newman-Keuls posttest turned up no significant differences between any one pair of means ($p > .05$). The effect of comparison series also appeared small (Table 1, top section). The analysis of variance showed that the difference in error magnitude on ascending and descending comparison judgments did not approach significance [$F(1, 36) = .84$].

A somewhat different picture of the data is obtained if not only magnitude but direction of error is also taken into consideration. Figure 2 shows a distinct difference between intra- and crossmodal judgments. Intramodal judgments generally produced a more random (even) distribution of positive errors (judgments larger than the standard) and negative errors (judgments smaller than the standard) than did crossmodal judgments. Crossmodal judgments revealed certain biases. The visual-haptic group produced generally positive errors, while the haptic-visual group generally produced negative errors. An overriding tendency to overset the most acute angle and underset the most obtuse angles also predominated, except in the visual-visual condition. The mean directional errors (in degrees) for ascending and descending conditions considered separately, and for overall bias, were -5.82, $+3.48$, and -1.17 for the haptic-haptic group; -2.72, $+2.44$, and $-.14$

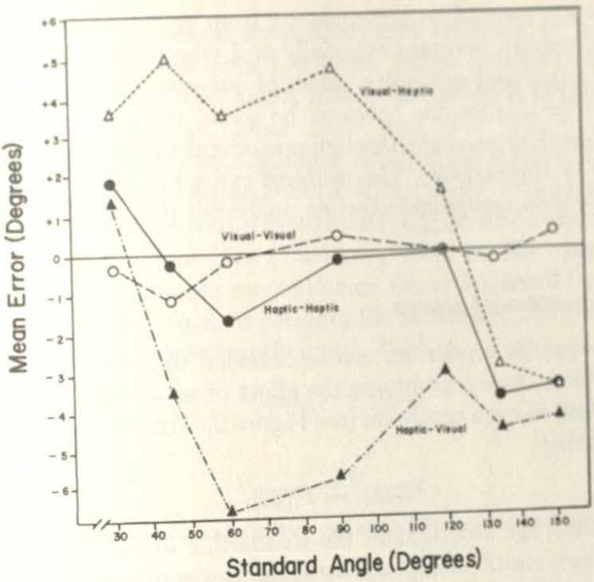


Fig. 2. Magnitude and direction of error (ascending and descending series combined)

for the visual-visual group; −7.61, +.20, and −3.70 for the haptic-visual group; and −3.64, +6.88, and +1.62 for the visual-haptic group.

DISCUSSION

In matching experiments, judgments of extent have traditionally been interpreted in terms of 'underestimation' and 'overestimation.' These interpretations have been based on the relationship between the true values of the standard and the comparison set to it (see Thurstone, 1948). For example, an acute-angle standard 'matched' by a 90-deg comparison would be interpreted as 'overestimated.' The implication is that the standard angle was perceived as larger (more obtuse) than it really was. However, this kind of interpretation is valid only where the true (objective) value of the comparison object is known to be a true measure of the subjective value of the standard object. In fact, this can never be assumed on an a priori basis. A 'match' defines a relative equality between the subjective value of the standard and the subjective (not objective) value of the chosen comparison.

The problem created by extrapolating from an actual value to a subjective relationship becomes especially real where the relationship between the actual and subjective values of one object is likely to be different from the relationship between the actual and subjective values of a paired object (for example, through procedural variables or crossmodal techniques of comparison). The problem exists for single features, but it is considerably compounded when analyzing perception of complex forms.

Intra- and crossmodal groups

Modality-specific tendencies can be revealed by examining the overriding algebraic error (combining the effect of ascending and descending movement) in each condition (see Figure 2). For all matches, it must be assumed that

$$X_{s(\mathrm{subj})} \approx Y_{c(\mathrm{subj})}\,, \tag{1}$$

where the subscript s designates the standard, c defines the final comparison setting, and X and Y represent the inspecting modalities. That is, in any match, the subjective value of the standard is approximately equal to the subjective value of the selected comparison.

For the *visual-visual* condition, we substitute in the general Equation 1

$$V_{s(\mathrm{subj})} \approx V_{c(\mathrm{subj})}\,, \tag{2}$$

where V is the notation for a visually inspected stimulus. Now, as Figure 2 shows, visual-visual errors have almost no overriding directional bias, and the magnitude of error is only a fraction of a degree. The equation expressing these facts is

$$V_{s(\mathrm{true})} \approx V_{c(\mathrm{true})}\,, \tag{3}$$

where $V_{s(\mathrm{true})}$ and $V_{c(\mathrm{true})}$ are the physical dimensions of the standard and comparison settings respectively. But Equations 2 and 3 can tell us nothing about the relationship between any true and subjective value. That is, correct comparison settings do not necessarily imply veridical percepts. They only imply an identical bias in the perception of standard and comparison. The magnitude of this bias could be any positive or negative number, or zero. Since matching experiments reveal only relative biases (standard versus comparison), other methods must be used to determine any absolute bias inherent in visual judgments of angle.

Accordingly, 20 additional subjects were asked to give verbal estimates of 'acute,' 'right,' or 'obtuse' to visually presented angles of 88, 90, and 92 deg. They showed no predominant tendency to perceive the angles as smaller (more acute) or larger (more obtuse) than true size. Presented with the 88-deg angle, 15% of the subjects judged it a right angle, 0% judged it an obtuse angle, and 85% judged it an acute angle; presented with a 90-deg angle, 70% judged it right, 15% obtuse, and 15% acute; and presented with a 92-deg angle, 15% judged it right, 85% obtuse, and 0% acute. These results suggest that visual percepts of angle have no inherent bias (Calabresi, 1933, on the other hand, reported that 90-deg angles seemed acute when inspected haptically). It seems justified, therefore, to consider the subjective and true values of visual judgments approximately equal. That is,

$$V_{s(\mathrm{subj})} \approx V_{s(\mathrm{true})} \tag{4}$$

and conversely,

$$V_{c(\mathrm{subj})} \approx V_{c(\mathrm{true})} \,. \tag{5}$$

For the *haptic-visual* condition, we substitute in Equation 1

$$H_{s(\mathrm{subj})} \approx V_{c(\mathrm{subj})} \,, \tag{6}$$

where *H* indicates a haptically inspected stimulus. As was just shown, the subjective value of a visual comparison is approximately equal to the true value of that setting (Equation 5). Combining Equations 6 and 5, therefore, we have

$$H_{s(\mathrm{subj})} \approx V_{c(\mathrm{true})} \,. \tag{7}$$

In words, the visually set comparison accurately predicts the subjective value of the haptic standard. Since this value is consistently lower than the true value of the standard (Figure 2), we are given the equation

$$V_{c(\mathrm{true})} < H_{s(\mathrm{true})} \,. \tag{8}$$

By combining Equations 8 and 7 it can be shown that

$$H_{s(\mathrm{subj})} < H_{s(\mathrm{true})} \,. \tag{9}$$

That is, relative to visual percepts, the haptic standard is perceptually smaller than its objective size. This supports the finding of Calabresi (1933).

If visual judgments are accurate predictors of objective size, and haptic judgments of angle have a relatively negative bias, what would be the effect if comparison judgments were haptic and standards were visual? We begin by substituting in the general Equation 1, for the *visual-haptic* condition

$$V_{s(\mathrm{subj})} \approx H_{c(\mathrm{subj})} \,. \tag{10}$$

But it was previously shown that the subjective value of a visual standard is approximately equal to its true value (Equation 4). Combining Equation 4 with Equation 10 gives us

$$V_{s(\mathrm{true})} \approx H_{c(\mathrm{subj})} . \qquad [11]$$

The data in Figure 2 show that in the visual-haptic condition the comparison settings are typically overset. That is,

$$V_{s(\mathrm{true})} < H_{c(\mathrm{true})} . \qquad [12]$$

Since the subjective value of the haptic comparison must be about equal to the subjective value of the visual standard (Equation 10), it must follow that the true values of the haptic comparisons are greater than their subjective values. This can be shown by combining Equations 11 and 12 so that

$$H_{c(\mathrm{subj})} < H_{c(\mathrm{true})} . \qquad [13]$$

Again, haptic percepts are relatively smaller than haptic stimuli.

The final case to consider is that of when both standard and comparison angles are inspected haptically—the *haptic-haptic* condition. It would seem logical that standard angles would share an identical bias with comparison angles (both would be subjectively smaller than their actual values) and that comparison settings would be approximately equal to the true value of the standards. However, a small but consistent pattern of undersetting was produced (Figure 2). That is,

$$H_{c(\mathrm{true})} < H_{s(\mathrm{true})} . \qquad [14]$$

This must mean that the bias inherent in perception of the standards is different from the bias inherent in perception of the comparisons. It was previously shown that the subjective values of both comparison and standard angles are smaller (when inspected haptically) than their objective size (Equations 12 and 13). And it is known from the general Equation 1 that these subjective values must be approximately equal to each other. Since the true value of the comparison setting is less than the true value of the standard (Equation 14), the difference between the subjective and true values of the standards must be greater than the difference between the subjective and true values of the comparisons. The magnitude of the negative bias, therefore, is greater for the standard than its comparison. These results are suggestive of a 'time error,' and implicate the role of haptic memory. It is possible that haptic bias increases within certain temporal parameters (or conversely, that subjective magnitude decreases with time). This could also explain why the PSEs were consistently lowest for the haptic-visual group.

Error has been shown typically to reflect the haptic rather than visual judgment. While visual percepts remain close to objective size, haptic percepts tend to be relatively smaller. When the standard angle is visual and the comparison is haptic (as it was for the visual-haptic group), the comparison must be overset to 'feel' as large as the standard 'looks.' When the standard is haptic and the comparison is visual (the haptic-visual group), the comparison must be underset, since it accurately matches the subjective value of the standard. Angularity is just one aspect of form, but an analysis of this kind can contribute to an understanding of modality-specific differences in overall perception of form as well. Such analyses should be applied when interpreting matching-experiment data.

Direction of angle movement

Some variables act only on the standard, some only on the comparison. Some variables act on both objects but in different ways. In the present study, direction of stimulus movement operated only on perception of the comparison object. Comparison stimuli presented in an ascending order (increasing angle size) were consistently set lower than when presented in a descending order (decreasing angle size). This means that ascending angles were judged relatively larger than stationary angles, while descending angles were judged to be relatively smaller than stationary angles. This tendency was large for touch, but vision was not immune. The effect for vision was small, but nevertheless consistent. This illustrates how a particular stimulus variable can influence matching judgments by affecting only a particular test object (standard or comparison) and by having a differential effect on two inspecting modalities. Procedural and modal factors can thereby influence the subjective values of specific features of form. Since this effect may vary among features or dimensions, the effect is multiplied when an entire form is being considered.

Calabresi (1933) tried to account for similar results in terms of the overt movement involved in setting the comparison. She felt that the amount of movement made in setting a haptic angle to 90 deg was usually overestimated. In the present experiment, however, vision was also susceptible to the influence of movement, although to a lesser degree. The under- and oversetting of angle does not seem to be strictly a property of overt motor activity. Rather, it seems necessary only that the value of the stimulus change in a particular direction. An angle subtending a continuously decreasing or increasing arc creates, respectively, the percepts of a smaller and larger arc than does a stationary stimulus. This phenomenon can be interpreted in terms of the psychophysical 'error of anticipation' or by adaptation-level theory (Helson, 1947). The latter would predict 'equal' judgments shifting toward the midpoint of the comparison series, that is, between the true value of the standard and the starting value of the comparison.

Notes

This paper is based on a thesis submitted to George Washington University in partial fulfillment of the requirements for the master's degree. The author expresses appreciation to Drs. R. D. Walk and J. J. Goodnow for their valuable advice. The study was supported in part by Grant R01 HD03105 to Dr. Goodnow from the National Institute for Child Health and Development. Received for publication February 3, 1971.

References

Appelle, S., and Goodnow, J. J. 1970. Haptic and visual perception of proportion. *Journal of Experimental Psychology* 84:47–52.

Behrman, B. W., and Brown, D. R. 1968. Multidimensional scaling of form: A psychophysical analysis. *Perception and Psychophysics* 4:19–25.

Björkman, M. 1967. Relations between intra-modal and cross-modal matching. *Scandinavian Journal of Psychology* 8:65–76.

Brown, D. R., and Andrews, M. H. 1968. Visual form discrimination: Multidimensional analysis. *Perception and Psychophysics* 3:401–406.

Brown, D. R., and Brumaghim, S. H. 1968. Perceptual equivalence, pattern perception, and multidimensional methods. *Perception and Psychophysics* 4:253–256.

Brumaghim, S. H. 1967. The equivalence of visually and tactually acquired information in the perception of form. *Dissertation Abstracts* 28:2637B.

Brumaghim, S. H., and Brown, D. R. 1968. Perceptual equivalence between visual and tactual pattern perception: An anchoring study. *Perception and Psychophysics* 4:175–179.

Calabresi, R. 1933. Sulla valutazione tattilo-cinetica degli angoli. *Archivio Italiano di Psicologia* 11:1–20.

Cashdan, S. 1968. Visual and haptic form discrimination under conditions of successive stimulation. *Journal of Experimental Psychology* 76:215–218.

Fisher, G. H. 1969. An experimental study of angular subtension. *Quarterly Journal of Experimental Psychology* 21(4):356–366.

Goodnow, J. J. 1969. Eye and hand: Differential sampling of form and orientation properties. *Neuropsychologia* 7:365–373.

Goodnow, J. J. 1970. Exploratory behavior: Relationships between what you find and how you find it. Paper presented at the Symposium of the Institute for Child Psychology, University of Minnesota, April.

Helson, H. 1947. Adaptation-level as a frame of reference for prediction of psychophysical data. *American Journal of Psychology* 60:1–29.

Jastrow, J. 1893. On the judgment of angles and position of lines. *American Journal of Psychology* 5:214–217.

Lederman, S. J., and Taylor, M. M. 1969. Perception of interpolated position and orientation by vision and active touch. *Perception and Psychophysics* 6(3):153–159.

Logan, O. 1964. Estimation and reproduction of angles from a given line of reference. *Perceptual and Motor Skills* 18:231–234.

Pratt, M. B. 1926. The visual estimation of angles. *Journal of Experimental Psychology* 9:132–140.

Revesz, G. 1950. *Psychology and art of the blind*. London: Longmans, Green.

Stenson, H. H. 1966. The physical factor of structure of random forms and their judged complexity. *Perception and Psychophysics* 1:303–310.

Thurstone, L. L. 1948. Psychophysical methods. In *Methods of psychology*, ed. T. G. Andrews. New York: Wiley.

Walk, R. D. 1965. Tactual and visual learning of forms differing in degree of symmetry. *Psychonomic Science* 2:93–94.

continued from page 472

Koch, Richard, and Dobson, James C., eds. *The Mentally Retarded Child and His Family*. New York: Brunner/Mazel, 1971. Pp. xi, 504. $15.00.

Krohn, Roger C. *The Social Shaping of Science*. Westport, Conn.: Greenwood Publishing, 1970. Pp. xxii, 280. $11.50.

Lake, Dale G. *Perceiving and Behaving*. New York: Teachers College Press, 1970. Pp. xiv, 102. $8.50.

Lazarus, Richard S., ed. *Personality*. 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xiv, 203. $5.75; paperback, $2.75.

Lee, S. G. M., and Herbert, Martin, eds. *Freud and Psychology*. Baltimore: Penguin Books, 1970. Pp. 398. Paperback, $1.75.

Levin, Harry, and Williams, Joanna P., eds. *Basic Studies on Reading*. New York: Basic Books, 1970. Pp. x, 283. $10.00.

McCarty, Donald J., and Ramsey, Charles E. *The School Managers: Power and Conflict in American Public Education*. Westport, Conn.: Greenwood Publishing, 1971. Pp. xxii, 281. $11.50.

McLaughlin, Terence. *Music and Communication*. New York: St. Martin's Press, 1971. Pp. 112. $6.95.

Maggio, Elio. *Psychophysiology of Learning and Memory*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xiii, 180. $10.00.

Marin, Peter, and Cohen, Allan Y. *Understanding Drug Use: An Adult's Guide to Drugs and the Young*. New York: Harper & Row, 1971. Pp. 157. $5.95.

Morgan, Clifford T., and King, Richard A. *Introduction to Psychology*. 4th ed. New York: McGraw-Hill, 1971. Pp. xiii, 784. $10.95.

Noland, Robert L. *Research and Report Writing in the Behavioral Sciences*. Springfield, Ill.: Charles C Thomas, 1970. Pp. viii, 98. $6.50.

Olson, David R. *Cognitive Development: The Child's Acquisition of Diagonality*. New York: Academic Press, 1970. Pp. xvi, 213.

Pailhous, Jean. *La représentation de l'espace urbain: L'exemple du chauffeur de taxi*. Paris: Presses Universitaires de France, 1970.

Perin, Constance. *With Man in Mind: An Interdisciplinary Prospectus for Environmental Design*. Cambridge, Mass.: The M.I.T. Press, 1970. Pp. 183. $7.50.

Pfeiffer, Wolfgang M. *Transkulturelle Psychiatrie: Ergenbisse und Probleme*. New York: Intercontinental Medical Book, 1971, Pp. x, 166.

Porter, Judith D. R. *Black Child, White Child: The Development of Racial Attitudes*. Cambridge, Mass.: Harvard University Press, 1971. Pp. xi, 278. $8.95.

Rank, Otto. *The Double: A Psychoanalytic Study*. Chapel Hill: University of North Carolina Press, 1971. Pp. xxii, 88. $5.00.

Reagles, Kenneth W.; Wright, George N.; and Butler, Alfred J. *A Scale of Rehabilitation Gain for Clients of an Expanded Vocational Rehabilitation Program*. Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1970. Pp. xvi, 74.

concluded on page 542

American Journal of Psychology
1971, Vol. 84, No. 4

SALIVARY HABITUATION: QUANTITATIVE SIMILARITIES TO SENSORY ADAPTATION

Charles H. Webb and Donald H. McBurney
University of Pittsburgh

The salivary responses of human subjects to acetic-acid stimulation increased over the first few trials each day, and on the first trial over days until the third day. Response decrement with repeated stimulation became more pronounced over days and was maximal by the third day. The magnitude of the salivary response increased as a logarithmic function of stimulus intensity, and the response to a strong stimulus was not diminished by prior habituation to a weaker stimulus.

Although adaptation and habituation are qualitatively similar phenomena, they have been distinguished traditionally on the basis of presumed differences in mechanism, locus, and time course (Thompson and Spencer, 1966). While these distinctions might once have seemed forceful, they are less convincing today (Martin, 1964). Considerable data suggest, for example, that mechanisms of sensory adaptation may not be confined to the receptors but may contain central components as well (Bartlett, 1965; Livingston, 1959).

With regard to time course, for another example, adaptation and recovery from adaptation have been thought to be short-term processes (seconds or minutes), and habituation and recovery from habituation have been shown to occur over longer periods of time (hours or days). Humphrey (1930, 1933), however, reported complete habituation of some responses in minutes or seconds, and Prosser and Hunter (1936) reported that recovery of a habituated response can occur within minutes. Furthermore, Hahn (1966) reported relatively long time courses (30 min) for vibrotactile adaptation; and as is well known, recovery from light adaptation can still be occurring after as long as 30 min of darkness. In fact, it has been reported (Clark, Johnson, and Dreher, 1946) that dark adaptation after repeated exposure to bright sunlight can take many hours, or even days.

In view of the difficulties of attempting to distinguish adaptation and habituation in the usual ways, it might be more fruitful to consider adaptation and habituation as two examples of a single fundamental behavioral process. A revaluation of this sort could be accomplished by demonstrating operational equivalences between the two types of response decrement. A number of such equivalences between adaptation and habituation are well known: both proceed as a negatively accelerated function of time, both are reversible, both exhibit a longer time course with higher stimulus intensities, and so on. (See Osgood, 1953, Thompson and Spencer, 1966, for parametric characteristics of adaptation and habituation respectively.) More rigorously quantitative equivalences need to be investigated, however, if adaptation and habituation are to be related in a convincing way.

For audition (Hedgecock, 1955), vision (Stevens and Stevens, 1963), taste (McBurney, 1966), and vibration (Gescheider and Wright, 1968), it has been shown that after adaptation to stimuli of higher intensities, the slope of the magnitude function (response magnitude as a function of stimulus intensity) becomes steeper than it was after adaptation to lower intensities (or with no adaptation at all), while the maximum response in both cases may remain the same. This steepening after auditory adaptation has been labeled recruitment. The present study was performed in order to determine whether steepening of the response function occurs after habituation. If it did, one could argue more forcefully for the behavioral similarity of adaptation and habituation.

A graded response is needed for such an experiment, since response magnitude must be measured on each trial. The salivary response to taste stimuli seemed appropriate. Since there is no report of salivary habituation in the literature, a pilot study was performed in order to demonstrate its occurrence. The results of the pilot experiment indicated that this response does habituate. Since this contradicts the findings of other investigators (Lashley, 1916; Kerr, 1961; Chauncey, Shannon, and Feller, 1967), a number of parametric characteristics of salivary habituation other than steepening were also studied.

METHOD

—Subjects—Forty male undergraduates at the University of Pittsburgh were selected on the basis of a pretest for salivation to acetic acid.

—Apparatus and solutions—Molar (M) solutions were made with distilled water and maintained at room temperature. Three solutions of acetic acid—.020 M, .064 M, and .20 M, one-half a log step apart in concentration—were used. Solutions were delivered from a 5-ml pipet. Saliva was collected in test tubes.

—Procedure—Thirty of the subjects were assigned randomly to three groups of 10 subjects each. Each of these subjects participated in the experiment for five days, randomly assigned to one of the three concentrations .020 M, .064 M, or .20 M of acetic acid, which he received for 10 habituation trials on each of the first three days in order to establish baseline rates of salivation and to determine whether warm-up effects were occurring. At the beginning of each trial the subject was given 2 ml of one of the solutions by pipet. He was told to draw the solution into his mouth and to sit with his head tilted downward and his tongue in the solution in his mouth for 1 min. At the end of a trial he was told to spit the contents of his mouth into a test tube. He then rinsed his mouth with distilled water and rested until the next trial. There was a 1-min interval between trials. The amount of saliva secreted during a trial was determined by measuring the contents of each test tube, then subtracting 2 ml.

Several other methods of collecting saliva were tried, including parotid caps. Since parotid salivation results primarily from stimulation of the posterior tongue, and since the convenient methods of taste stimulation result primarily in stimulation of the anterior tongue, it was found that the collection of whole saliva would yield more consistent results than the collection of only parotid saliva.

After day 3 the 30 subjects were again randomly assigned to one of three groups (I, II, and III) for days 4 and 5. The treatment plan is shown in Table 1. On day 4 the subjects were given 10 habituation trials with the stimulus to be used on the test trials of day 5. On day 5 they were given a weaker stimulus for 10 habituation trials followed by 2 test trials with a stronger stimulus. A comparison of (*a*) the size of the response decrement between trials 1–2 and trials 9–10 on day 5 with (*b*) the difference between the first two responses on day 4 and the test-trial responses on day 5 constituted the test for steepening.

Table 1. Treatment plan for steepening

Group	Day 4 Habituation trials	Day 5 Habituation trials	Day 5 Test trials
I	.20 M	.020 M	.20 M
II	.20 M	.064 M	.20 M
III	.064 M	.020 M	.064 M
IV	.064 M	.20 M	.064 M

The 10 additional subjects (group IV) were selected after the first phase of the experiment, described above, and participated in a *control* experiment. They were also in their experiment for five days. On days 1 and 2 all these control subjects were given 10 habituation trials with .064-M acetic acid in the manner already described. On day 2 five of these subjects who had habituated were exposed to a strong dishabituating stimulus (a bell accompanied by the clatter of a bottle dropped on the floor) in the interval separating trials 8 and 9. On day 3 half of the 10 subjects, randomly chosen, were given 10 habituation trials with .064-M acetic acid; these trials were separated by 2-min intervals rather than the 1-min intervals used before. The other half of the subjects were given 10 habituation trials with the same concentration, but their trials were separated by 15-sec intervals. On day 4 the 2-min and 15-sec intertrial intervals were reversed, and the subjects were again given 10 trials with the .064-M solution. On day 5 all of these control subjects got 10 habituation trials with .20-M acetic acid, followed by 2 test trials with .064-M acetic acid.

RESULTS

Warm-up

The arithmetic mean amount salivated on each trial was calculated for each day and concentration for those 17 (of the first 30) subjects who received the same concentration on days 1, 2, and 4 or 5. Five subjects got .020-M, seven got .064-M, and five got .20-M solutions. Figure 1 shows the mean amount salivated as a function of trials, days, and concentration for these subjects.

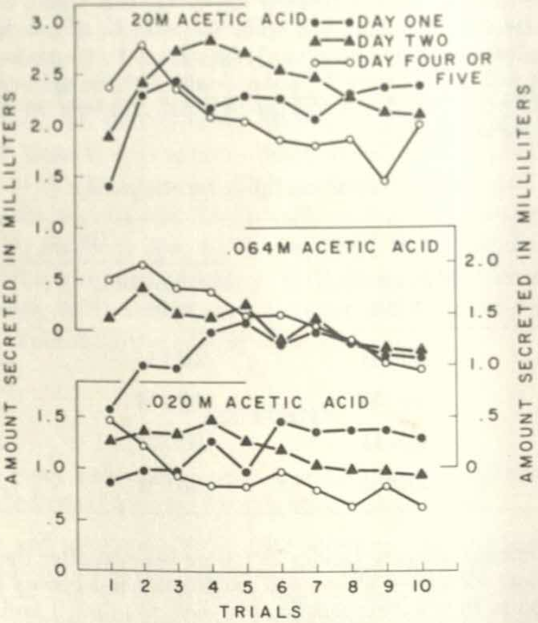


Fig. 1. Warm-up effects and habituation of the salivary reflex as a function of trials and days with three stimulus concentrations of acetic acid

Amount salivated per trial on day 1, day 2, and day 4 or 5 was averaged across concentrations. A test of the linear trend as a function of trials was performed. There was a warm-up effect on day 1 that was significant across concentrations [$t(16) = 3.73, p < .01$]. Amount salivated on trial 1 increased from day 1 to day 4 or 5 [$t(16) = 7.22, p < .001$]. Although it is not shown in Figure 1, it should be noted that salivation on trial 1 of days 4 or 5 was not greater than salivation on trial 1 of day 3, indicating that the day–warm-up effect did not proceed beyond day 3.

Response as a function of concentration

The upper line in the composite data of Figure 3 shows mean salivation rate before habituation as a function of stimulus concentration. Each point is mean salivation on the first two trials of a sequence for groups I–III on day 4 or 5 of the experiment. Salivation rate increased as a logarithmic function of stimulus concentration.

Habituation

On day 2 there was a habituation trend that was significant across concentrations [$t(16) = -3.49, p < .01$]. This trend became more pronounced by day 4 or 5 [$t(16) = -9.20, p < .001$]. Salivation on trial 10 decreased between day 1 and day 4 or 5 [$t(16) = 3.71, p < .01$]. On day 4 or 5 there was a 50.4% decrease in response magnitude over the 10 habituation trials when the .020-M solution was the stimulus [$t(4) = -6.97, p < .01$], a 44.4% decrease when the .064-M solution was the stimulus [$t(6) = -10.60, p < .001$], and a 45.2% decrease when the .20-M solution was the stimulus [$t(4) = -3.74, p < .02$]. No habituation was found across days.

One-minute intervals separated the habituation trials shown in Figure 1. The top part of Figure 2 shows average salivation rate per trial to .064-M acetic acid for the subjects in group IV, who had 2-min and 15-sec intervals between trials on days 3 and 4. Data from the 10 subjects in group IV on those days are represented in each curve. There was no significant difference between rates of habituation with the different intertrial intervals.

Five subjects in group IV who had shown habituation on day 2 were tested for dishabituation by presenting a loud extraneous stimulus in the interval between trials 8 and 9. The bottom part of Figure 2 shows that a 63.5% increase in response occurred [$t(4) = 12.66, p < .005$].

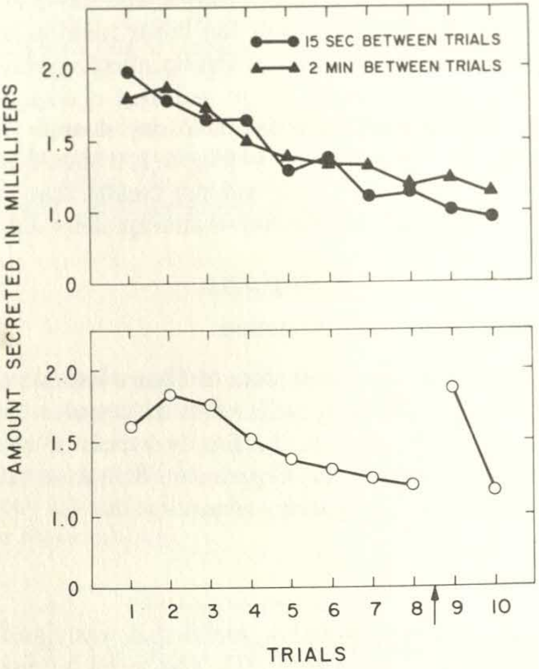


Fig. 2. Upper panel, Habituation of the salivary reflex to .064-M acetic acid with two frequencies of stimulus presentation. Lower panel, Dishabituation of the salivary reflex following habituation to .064-M acetic acid. (A strong extraneous stimulus was presented, as shown by the arrow, between trials 8 and 9.)

Steepening

Before- and after-habituation responses to habituating and test stimuli are plotted for groups I, II, III, and IV in Figure 3. For each group the plotted magnitude of the response to the habituating stimulus before habituation is the mean of the first two trials with that stimulus on day 5 of the experiment, and the response to the habituating stimulus after habituation is the mean of the last two trials on day 5. The response to the test stimulus before habituation is from data for the first two trials on day 4 for each group, and the response to the test stimulus after habituation is the average on the two test trials presented after habituation on day 5 for each group. (Since no habituation occurred over days, the response to the stimulus on day 5 can be compared to the response to the test stimulus presented on day 5; see Table 1. Note also that the responses to the test stimulus are plotted at the right of the figures for groups I, II, and III, each of which had a test stimulus stronger than the habituating stimulus; and at the left of the figure for group IV, which had a test stimulus weaker than the habituating stimulus.)

In every case that a stronger—test—stimulus was presented after habituation to a weaker stimulus (see Figure 3 for groups I, II, and III), the response to the stronger stimulus was either as large or almost as large after habituation as it was before habituation. In none of these groups was the response to the stronger stimulus diminished as much after habituation as was the response to the habituating stimulus itself. But when a weaker—test—stimulus was presented after habituation to a stronger stimulus (see Figure 3 for group IV), the response to weaker stimulus was not as great after as before habituation.

In order to test the significance of the steepening effect, the data of the three groups that received a stronger concentration after habituation (groups I, II, and III) were pooled over the specific concentrations that constituted the weaker and stronger conditions. A 2 × 2 within-subjects analysis of variance was performed. For each subject in groups I–III there were four measures (see Figure 3): before and after habituation, and weaker and stronger concentrations. In this design the steepening effect is represented by the interaction. A significant interaction was obtained [$F(1, 116) = 15.41, p < .001$]. When a similar analysis was performed on the data from group IV the interaction was not significant [$F(1, 36) = .04$]. The bottom panel of Figure 3 is a composite representation of the steepening effect derived from the data of groups I, II, and III.

DISCUSSION

As already mentioned, other investigators (Lashley, 1916; Kerr, 1961; Chauncey, Shannon, and Feller, 1967) have reported that habituation of the salivary reflex does not occur. Yet the data in Figures 1 and 2 indicate that habituation of the salivary response in humans did occur with each of the three stimulus concentrations used. (Ten habituation trials did not produce complete cessation of the response, but complete habituation of salivation would not be expected.) It is possible that the failure of other investigators to demonstrate salivary habituation is due, in part, to the fact that they tested their subjects on only one day.

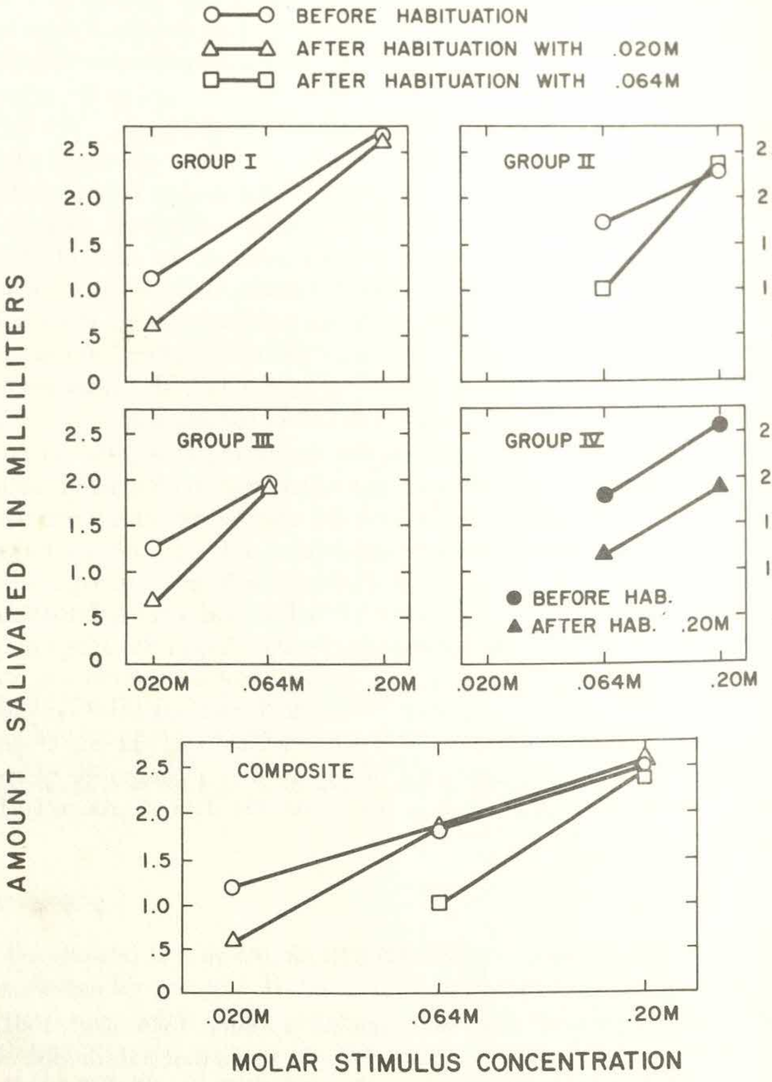


Fig. 3. Steepening; salivary responses to habituating and test stimuli before and after habituation. (In groups I–III, and in the composite data for them, the test stimulus was stronger than the habituating stimulus and the response to the test stimulus was not diminished by habituation; steepening did occur. In group IV the test stimulus was weaker than the habituating stimulus, and steepening did not occur.)

Habituation did not occur in most subjects on the first day of the present experiment either. Figure 1 indicates that response magnitude on early trials increased rather than decreased on day 1, and also increased between day 1 and day 4 or 5. Habituation occurred to some extent on day 2 and was pronounced by day 4 or 5, when maximum responding had been achieved. Had this warm-up effect been taken into account in other experiments, habituation would probably have been observed.

The bottom part of Figure 2 indicates that dishabituation of the response occurred, as expected, when a loud extraneous stimulus was introduced toward the end of a 10-trial sequence. The dishabituating stimulus did not seem to cancel the habituation completely, however, since the response returned to its habituated level 2 trials after the dishabituating stimulus was presented. This result lends support to the contention of Thompson and Spencer (1966) that dishabituation is a sensitization process independent of the habituation process itself.

The results presented in the upper line of the composite data in Figure 3 indicate that salivary flow rate increased as a function of stimulus concentration, a finding consistent with the findings of other investigators (Kerr, 1961; Chauncey, Shannon, and Feller, 1967). The relationship between stimulus concentration and response magnitude seems to be logarithmic. Kerr (1961) discusses a considerable body of additional evidence that supports this conclusion.

The results presented in Figure 3 indicate that steepening of the response function occurred after salivary habituation when the test stimulus was stronger than the habituating stimulus. This finding is analogous to the finding in the sensory context, already discussed, that the sensory magnitude of a strong stimulus is not diminished by adaptation to a weaker stimulus of the same modality and quality (Hedgecock, 1955; Stevens and Stevens, 1963; McBurney, 1966; Gescheider and Wright, 1968). As might also have been expected, steepening did not occur after salivary habituation when the test stimulus was weaker than the habituating stimulus. Bernstein (1968) has recently reported similar failure in response recovery to weaker stimuli after habituation of the orienting response.

The occurrence of steepening in habituation has been implied before (Sherrington, 1906, on spinal 'scratch' reflexes in the dog; Humphrey, 1933, on habituation of leg-withdrawal responses in the musk turtle; Sharpless and Jasper, 1956, on habituation of electroencephalographic arousal; Voronin and Sokolov, 1960, on habituation of the orienting response), but it has never been studied in a systematic or quantitative way. Frequently, steepening and dishabituation have been confused.

Thompson and Spencer (1966), for example, mention the finding by Voronin and Sokolov (1960) that when the intensity of an auditory stimulus to which subjects were habituated was decreased, the orienting response to the weaker stimulus recovered. Thompson and Spencer consider this an example (albeit a special case; 1966, p. 33) of dishabituation, when actually it is an example of steepening similar to that found in the *sensory* context and mentioned above.

Operationally, it must be noted, steepening entails habituation (or adaptation) with a stimulus of one intensity, followed by presentation of either a stronger or weaker test stimulus of the same modality and quality. The response to the test stimulus is used as the measure of steepening. This paradigm is equivalent to an intensity-generalization paradigm, with failure of generalization being analogous to the occurrence of steepening. Dishabituation, on the other hand, entails habituation with a stimulus of some intensity, after which an extraneous stimulus, usually very strong and often involving a different modality than the habituating stimulus, is presented. Then the original habituating stimulus is presented again, and it is the response to *this* stimulus (rather than to the second stimulus, which is analogous to the test stimulus in the steepening paradigm) that is used as the measure of dishabituation. Perhaps this confusion has prevented the inclusion of steepening as a parametric characteristic of habituation in other accounts.

The quantitative demonstration of steepening after habituation is important because it indicates another behavioral equivalence between adaptation and habituation. The warm-up effect found in the present study is also important in this regard, since such an effect has been demonstrated in various sensory modalities as well as other types of responses (e.g., Franzisket, 1963, on leg-wiping responses to tactile stimulation in the frog).

The occurrence of such behavioral equivalences in seemingly different types of decremental processes suggests that these processes may not be as distinctive as had once been thought. This notion has been considered by various other authors (Gampel, 1966; Verplanck, 1966; Martin, 1964), although little research aimed at demonstrating it has been undertaken. However, Drever and Bower (1964) did explore the relationship between adaptation, satiation, and habituation in a series of visual-rearrangement studies, and Leaf, Kayser, Andrews, Adkins, and Leaf (1968) have investigated the relative efficacy of habituation and extinction in the blocking of fear conditioning. The results of both studies imply close relationships among the phenomena in question. The present study lends support to the proposition that adaptation and habituation are merely two artificially distinguished examples of the same behavioral process. When viewed in this light, adaptation may be considered a fundamental process parallel to, say, respiration or digestion—a process that serves the same function in every organism, although the mechanisms involved may differ between and even within organisms.

Notes

This research was supported by United States Public Health Service predoctoral fellowship 1-F1-MH-41,999-01 to the first author and by USPHS Grant NB 07873–03 to the second author. The authors thank William Ray for helpful suggestions. Received for publication September 10, 1970.

References

Bartlett, N. R. 1965. Dark adaptation and light adaptation. In *Vision and visual perception,* ed. C. H. Graham, pp. 185–207. New York: Wiley.

Bernstein, A. S. 1968. The orienting response and direction of stimulus change. *Psychonomic Science* 12:127–128.

Chauncey, H. H., Shannon, I. L., and Feller, R. P. 1967. Effect of oral and nasal chemoreception on parotid gland secretion. In *Secretory mechanisms of the salivary glands,* ed. L. H. Schneyer and C. A. Schneyer, p. 351. New York: Academic Press.

Clark, B., Johnson, M. L., and Dreher, R. E. 1946. The effect of sunlight on dark adaptation. *American Journal of Ophthamology* 29:828–836.

Drever, J., and Bower, T. G. R. 1964. Adaptation, satiation and habituation: A taxonomic study. *Acta Psychologica* 22:202–217.

Fisher, R. A., and Yates, F. 1943. *Statistical tables for biological, agricultural, and medical research.* Edinburgh: Oliver & Boyd.

Franzisket, L. 1963. Characteristics of instinctive behavior and learning in reflex activity of the frog. *Animal Behavior* 11:318–324.

Gampel, D. H. 1966. Data parallels across behavior systems. *Psychological Bulletin* 66:499–510.

Gescheider, G. A., and Wright, J. H. 1968. Effects of sensory adaptation on the form of the psychophysical magnitude function for cutaneous vibration. *Journal of Experimental Psychology* 77:308–313.

Hahn, J. F. 1966. Vibrotactile adaptation and recovery measured by two methods. *Journal of Experimental Psychology* 71:655–659.

Hedgecock, L. D. 1955. The measurement of auditory recruitment. *A.M.A. Archives of Otolaryngology* 62:515–527.

Humphrey, G. 1930. Le Chatelier's rule, and the problem of habituation and dehabituation in *Helix albolabris. Psychologische Forschung* 13:113–127.

Humphrey, G. 1933. *The nature of learning in its relation to the living system.* New York: Harcourt, Brace.

Kerr, A. C. 1961. *The physiological regulation of salivary secretion in man.* Oxford: Pergamon Press.

Lashley, L. S. 1916. Reflex secretion of the human parotid gland. *Journal of Experimental Psychology* 1:461–493.

Leaf, R. C., Kayser, R. J., Andrews, J. S., Adkins, J. W., and Leaf, S. R. 1968. Block of fear conditioning induced by habituation or extinction. *Psychonomic Science* 10:189–190.

Livingston, R. B. 1959. Central control of receptors and sensory transmission systems. In *Handbook of physiology*. Vol. 1, Section 1, *Neurophysiology*, ed. J. Field, H. W. Magoun, and V. E. Hall, pp. 741–760. Washington, D.C.: American Physiological Society.

Martin, I. 1964. Adaptation. *Psychological Bulletin* 61:35–44.

McBurney, D. H. 1966. Magnitude estimation of the taste of sodium chloride after adaptation to sodium chloride. *Journal of Experimental Psychology* 72:869–873.

Osgood, C. E. 1953. *Method and theory in experimental psychology*. New York: Oxford University Press.

Prosser, C. L., and Hunter, W. S. 1936. The extinction of startle responses and spinal reflexes in the white rat. *American Journal of Physiology* 117: 609–618.

Sharpless, S., and Jasper, H. 1956. Habituation of the arousal reaction. *Brain* 79:655–682.

Sherrington, C. 1906. *The integrative action of the nervous system*. New Haven, Conn.: Yale University Press.

Stevens, J. C., and Stevens, S. S. 1963. Brightness function: Effects of adaptation. *Journal of the Optical Society of America* 55:375–385.

Thompson, R. F., and Spencer, W. A. 1966. Habituation: A model phenomenon. *Psychological Review* 73:16–43.

Verplanck, W. S. 1966. The operational investigation of behavior. Unpublished outline, University of Tennessee.

Voronin, L. G., and Sokolov, E. N. 1960. Cortical mechanisms of the orienting reflex and its relation to the conditioned reflex. *Electroencephalography and Clinical Neurophysiology* suppl. 13:335–346.

American Journal of Psychology
1971, Vol. 84, No. 4

CONSTRUCTION AND RECONSTRUCTION IN MEMORY

Stanley A. Summers and James S. Fleming
San Fernando Valley State College

The effect of cues on retention of written material was compared to their effect on pure construction, without previous reading. Each subject answered questions on selections, some of which he had read 2, 7, or 28 days before, and some of which were partly available while he answered, to provide cues from context. Both context and retention significantly enhanced accuracy, but their effects were not additive; context added nothing to the accuracy from retention. Retentional processes may involve both direct reconstruction and an indirect construction that duplicates performance based on contextual cues.

In this experiment, contextual aids to correct judgment and interpretation were added to the other influences on memory performance. Accuracy resulting from these combined influences was compared to accuracy from the judgmental influences alone, and to accuracy from ordinary retention alone. The contextual aids provided a judgmental basis for accuracy, that is, for a process of construction. Also, any redintegrative effect, of enhancement of retention by the presented context, could be observed.

For many experimental questions it is not important to distinguish accuracy based on learning and retention from accuracy based on judgment and guess, and experimental operations often do not clearly ascribe memory performance to one component or another. However, the conceptual distinction between retentional and interpretational components has been important, especially for long-term memory of complex material, and retention itself sometimes has been considered as a construction process.

Consider, for example, the tradition of strong emphasis on the cognitive shaping of previous experience that was set by Kuhlman, who believed that “the total process of recall is very largely not recall at all, and can never be described even half correctly by calling it reproduction. It is rather a construction, not a reconstruction, a construction of a certain result that is accepted in place of the original, and far from a reconstruction of a past perception" (1906, pp. 342–343). The cognitive shaping of memory has also been emphasized by Bartlett (1932) and more recently by Jarvik (1966) and Neisser (1967). Generally, these writers dealt with the interplay between the material stored from previous experiences and the other determinants present at the time of remembering.

Subjects in the present experiment answered questions about selections of writing. In some conditions the subject had previously read the selection on which he was questioned. At times the bulk of the selection was available to him while he answered questions but the specific passages on which questions were based were blanked out. Any cuing of retention through the context provided by the remaining passages could be measured. In addition, the influence of judgment and interpretation based purely on context, without prior exposure to the missing material, could be measured. This study, then, analyzed the separate and joint effects, on accuracy, of present context and of previous exposure.

METHOD

—Subjects—The subjects were volunteers from undergraduate psychology classes. Prior to the experiment 20 subjects were assigned to a special condition for establishment of criteria of accuracy. Then 45 comparable subjects took part in the experiment proper.

—Preliminary establishment of criteria—The answers given to prospective questions by the group of 20 subjects, with the complete written material before them, provided the standard for accurate retention. These subjects read each selection and then answered questions on it, rereading and interpreting the parts relevant to each answer as they did so. Questions for the experiment were chosen from those on which at least 19 of the 20 subjects agreed with each other on the answer, as well as with the keyed answer previously chosen by the experimenters. A match to these answers, then, was considered perfect interpretation and retention.

—Materials—In the experiment proper, there were 16 selections, each of between 500 and 1,000 words. These selections were mainly narrative; half were selected from fiction and the rest from history, criticism, and general exposition. The content varied considerably in locale, era, and subject matter, and the selections differed in style and quality; they were distinct to the subjects. They were also usually interesting, and occasionally slightly familiar, to them.

Each selection was labeled by source and by author. The selections were presented on Xeroxed typed sheets, and each was reproduced both in a complete version and in a version with an average of approximately 10% of the typed material painted out before reproduction. The sentences painted out contained the specific information on which a question was based; the remaining material, approximately 90% of the original selection, contained the verbal context for this information. Lengths and placements of the blanked-out portions varied. The questions were multiple choice, with six alternatives each. Three questions for each selection appeared on a single page, making 16 pages in all, stapled together in different random orders.

The following example gives part of one of the selections, from *Walden* by Henry David Thoreau, and one question on it. The material within brackets, which was specific to the question, was blanked out in some conditions. Material outside the brackets provided the context for the information asked by the question.

> Standing on the snow-covered plain, as if in a pasture amid the hills, I cut my way first through a foot of snow, and then through a foot of ice, and open a window under my feet, where, kneeling to drink, I look down into the quiet parlour of the fishes, pervaded by a softened light as through a window of ground glass, with its bright sanded floor the same as in summer; there a perennial waveless serenity reigns as in the amber twilight sky, corresponding to the cool and even temperament of the inhabitants. Heaven is under our feet as well as over our heads.
>
> [Early in the morning, while all things are crisp with frost, men come with fishing reels and slender lunch, and let down their fine lines through the snowy field to take pickerel and perch;] wild men, who instinctively follow other fashions and trust other authorities than their townsmen, and by their goings and comings stitch towns together in parts where else they would be ripped.
>
> The men come to fish on the pond
> a. At high noon
> b. Early in the morning
> c. In the warmth of afternoon
> d. In evening twilight
> e. Just before dawn
> f. Arriving throughout the day

—Design and procedure—Each subject participated in the experiment on two different days. On the first day, he read 8 complete selections and then, either 2, 7, or 28 days later, answered questions on all 16 selections. Each time span between reading selections and answering questions was assigned to an experimental group of 15 subjects, and within these groups the subjects participated in smaller subgroups. In each session the subjects had as much time as they desired, but they usually finished in less than an hour. At the first session they were instructed:

> Please read thoroughly the eight selections you will be given today. You will have as much time as you need to absorb them well, a full hour if you want. Do not look up any of these materials later; just learn them as well as you can here. Next time you will be given more materials as well as some explanation of the purpose of the experiment, but for now we won't go into questions of what this is about.

At the next session (2, 7, or 28 days later), they were instructed:

> Today, you will be asked a series of multiple-choice questions concerning sixteen different narrative selections. Do as well as you can on all the questions, even though the material on which they are based will be more familiar for some questions than for others. You will remember some of the material from the selections you read before. For some of the selections on which the questions are based, you will have *part* of the narratives available when you answer the questions; please use them fully for the questions they apply to. Even though you may sometimes feel you don't know the answer at all, try hard on each question and, if you have to, make the best guess you can.

At this second session all the subjects answered the same 48 questions, 3 questions on each selection of writing. They received a selection under only one experimental condition, answering 12 questions under each of four experimental conditions. The selections appeared in nearly balanced frequency and combination under each condition. These conditions were based on whether the selection had been presented at the previous session and on whether the context material was presented with the questions. Each subject read half the selections in complete form at the first session, and each received context materials for half the selections at the second session. This 2 × 2 design was replicated, with the different experimental groups, over the three intervals between sessions.

Under the *retention-plus-context condition,* the subject had read the complete selection previously and had the context passages available for reference while answering questions. Under the *retention condition,* he had read the complete selection at the previous session but had no relevant reference materials available while answering questions. Under the *context condition,* he had not read the complete selection but did have the relevant context passages available for reference while answering questions. And under the *baseline condition,* he neither had read the selection nor had relevant reference materials available while answering questions. All the context material, for the retention-plus-context condition and the context condition, was equally available for reference, independent of whether the subject previously had read the complete selection. This material was presented in a single stack with the eight selections for a subject arranged in a random sequence; usually it was necessary for him to leaf through several of these selections to find any particular one.

RESULTS

The data were analyzed to determine whether the context and the retention conditions each would provide accuracy beyond that of the baseline condition. It also was asked whether context or retention provided greater accuracy and whether this relationship changed over time. Finally, it was asked how the effects of context and retention interacted and, specifically, whether context provided a redintegrative aid to retention. Figure 1 presents the separate effects for the four conditions of the 2 × 2 design. The relative position of the four conditions appears consistent, except for a decrease in the advantage of the two retention conditions over the two nonretention conditions.

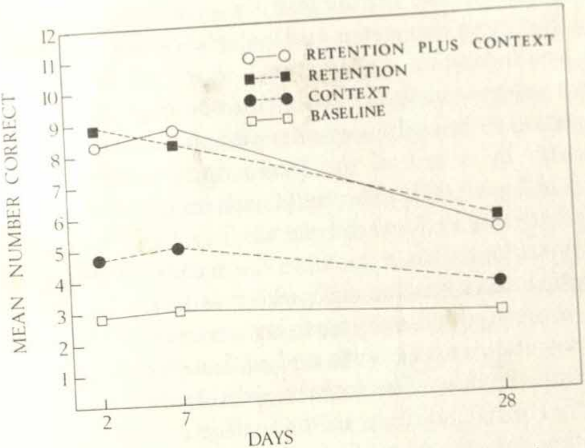


Fig. 1. Mean number correct per subject as a function of days since original presentation

An analysis of variance tested the overall effects of the experimental conditions on number of correct answers. The context effect was significant [$F(1, 42) = 10.9, p < .005$]. That is, the subjects answered significantly more questions correctly in the conditions in which they had context materials available for reference than in the conditions in which they did not. The retention effect was also significant [$F(1, 42) = 139.7$, $p < .001$]; the subjects correctly answered significantly more questions on material they had read at the first session, showing continuing effect from this exposure. As would be expected if the context provided a judgment situation independent of amount of days since the first session, there was no significant context × day effect [$F(2, 42) = 1.2$]. However, consistent with Figure 1, there was a significant retention × day effect [$F(2, 42) = 5.3, p < .01$], reflecting the decline of accuracy, due to forgetting, with days between presentation and test.

The overall proportions, across days, correct for each condition were: retention-plus-context, .63; retention, .63; context, .36; and baseline, .23. These proportions may be compared to an expected value of .17 for purely random responses. A t test of difference in number correct in the context and baseline conditions showed the performance advantage from context to be significant [$t(44) = 4.6, p < .001$]; the subjects did construct successfully from context. In turn, the retention condition had an advantage in number correct over the context condition [$t(44) = 7.8$, $p < .001$]. For these time spans and materials at least, the reconstruction (and any accompanying construction) provided through previous exposure was superior to the construction made possible by the context.

The question of how the experimental components combined is answered partly by a test of the interaction of context and retention [$F(1,42) = 12.1, p < .005$]. The significant interaction indicates that the context and retention effects did not combine additively; both Figure 1 and the proportions correct indicate the nature of this interaction. Although context and retention each added accuracy beyond the baseline condition, context added nothing to the accuracy available from retention. The retention curves, with and without context, were extremely similar over the 28 days. The context and retention conditions, effective separately, were not additive in their effects: context in combination with retention did not aid performance.

DISCUSSION

The circumstances that favor construction do not necessarily provide accurate construction. In this experiment, however, context did act as an alternative to previous exposure in providing accuracy. The context passages had implicit in them some part of the missing passages, a part substantial enough to enhance accuracy on unfamiliar material. Given this, it may seem surprising that such cuing from context provided no additional accuracy on materials previously viewed. The subjects went through their stacks of interleaved context materials, using them successfully in the context condition, but making no successful use of them in the retention-plus-context condition. There was no indication in their actions or comments that they selectively ignored context from passages they had read before.

These results can be conceptualized in terms of the memory processes already discussed. It has been assumed here that a large component of judgment and reasoned guess is elicited by many retention tasks. Such a process should have aided answers in both the context and the retention conditions and may, in itself, have provided comparable accuracy in these conditions. In the retention condition, this construction process would be an extrapolation from *stored* material; in the context condition, construction would be an extrapolation from the externally presented material. An additional basis for accuracy, available only in the retention condition, is the more direct activation and application of the stored material, when little interpretation and extrapolation are required for a correct answer. This aspect of memory accounts for the margin of advantage of the retention over the context condition; that is, both conditions involve similar construction processes, but only the retention condition involves reconstruction as well.

In the retention-plus-context condition, the stored material appears to have been sufficiently rich and detailed to provide the subjects with as much basis for judgment as they could use successfully, making redundant any additional aids from context. It can be inferred further that the specific cues available from the presented context may have been similar enough to cues already available from retained experience as to add little to performance. Given such similarity of cues, the response of the subjects to the physical form of the narrative, together with its gaps, would provide a close analogue of their response to the form in which they have stored material.

Viewed differently, this retention-plus-context condition did not produce any redintegrative effect; presentation of the bulk of the narrative did not evoke the previously experienced remainder. This contrasts with some of the results discussed by Horowitz and Prytulak (1969). However, their results mainly concerned short-term memory for brief materials, like paired associates. A broken-unit effect, which apparently aids redintegration, would be less prominent in longer-term memory or for more complex materials, although the concept of redintegration often has been invoked to deal with remote experiences and abstract relations.

For the 28-day span of the present study, the subjects' performance in the retention condition was superior to their performance in the (printed) context condition. To the extent that a *stored* context, retained from previous exposure, provided their basis for response, this stored context remained as useful as the printed one. If the retention curve of Figure 1 continued to decline at a constant rate and the context curve remained roughly level, the curves would intersect in a few months. At that point, memory performance would become more dependent on the cognitive use of printed context than on the cognitive use of stored context.

Notes

The authors thank Karyn Massoni for her valuable help in conducting the experiment. Received for publication October 26, 1970.

References

Bartlett, F. C. 1932. *Remembering*. New York: Macmillan.

Horowitz, L. M., and Prytulak, L. S. 1969. Redintegrative memory. *Psychological Review* 76:519–531.

Jarvik, M. E. 1966. A functional view of memory. In *The psychology of Egon Brunswik*, ed. K. R. Hammond. New York: Holt, Rinehart and Winston.

Kuhlman, F. 1906. On the analysis of the memory consciousness: A study in the mental imagery and memory of meaningless visual forms. *Psychological Review* 13:316–348.

Neisser, U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.

American Journal of Psychology
1971, Vol. 84, No. 4

FORMAL INTRALIST RESPONSE SIMILARITY: ITS ROLE IN PAIRED-ASSOCIATE LEARNING

John F. Hall
Pennsylvania State University

Two experiments using low-meaningfulness CVCs as responses were conducted in order to examine the role of formal intralist similarity on the response-learning and associative-learning stages that have been hypothesized to characterize the learning of paired associates. The results indicate that both stages were facilitated by low intralist similarity.

On the role of formal intralist response similarity in paired-associate learning, controversial findings have been reported. Goss and Nodine (1965) have reported that such similarity hinders learning, while Underwood (1953) has reported that it has no influence.

A more recent study by Jung (1965) is noteworthy inasmuch as this experimenter was interested in examining the role of formal intralist response similarity on both the response-learning and associative stages of paired-associate learning. Modeling his procedure after that employed by Underwood, Runquist, and Schulz (1959), Jung (1965) employed a free-recall test to measure the response-learning stage, and a matching test (here, all stimuli and responses that made up the conventional PA test were available and the subject had only to pair them correctly) to measure the associative or hook-up stage. In his first experiment, Jung found that intralist response similarity did not contribute to either stage; in his second experiment, his results indicated that low similarity resulted in associative learning being facilitated, but not response learning.

A basic difficulty with Jung's (1965) first experiment was that all subjects were given both the response-learning test and then the matching test on each of the five trials that he provided. As Jung has acknowledged, providing the subject with both tests undoubtedly resulted in some contamination of the findings, since "some response learning certainly may occur which will serve to inflate the response-recall test score of the next trial. Similarly, the execution of the response-recall test may affect [the subject's] performance on the associative-matching test which follows it" (1965, p. 377). In his second experiment, although each stimulus-response pair was presented, Jung instructed his subjects to learn either (*a*) just the response items (response learning) and not the pairings, or (*b*) the specific pairings (associative learning) with recall of the response items not being required. The problem here, of course, is that these instructions change the nature of the task, so that we cannot be sure that we are continuing to deal with the learning of paired associates as such.

Jung's (1965) procedures are in contrast to the purer experimental design found in the Underwood, Runquist, and Schulz (1959) study, which used separate groups to obtain the response-learning measure as well as separate groups after different numbers of PA trials. In view of Jung's methodological difficulties, an experiment using a design similar to that used by Underwood, Runquist, and Schulz (1959) was conducted, in order to simultaneously examine the role of formal intralist response similarity on both response and associative learning.

EXPERIMENT I

Method

—Materials—CVCs of high and low formal similarity were used to construct the responses for the two lists of material. The high-similarity responses consisted of these nine CVCs: CAH, CEG, CIQ, HAX, HEG, HIX, RAX, REZ, RIQ. The low-similarity responses were: HEQ, CIQ, DAH, FUB, LEH, RAX, JUZ, POB, and WOS. The two lists were equivalent with respect to Noble's (1961) rated association value (a'). Two-digit numbers were used as stimuli.

—Subjects—One hundred and eighty volunteers from an introductory psychology course were used as subjects. These were divided into 18 groups with 10 subjects per group. Nine groups learned each list. For each list, four groups were free-recall learning groups; here a single *free-recall trial* was provided each group at either trial 2, 5, 8, or 11. Four other groups were associative-learning groups. Here, a single *matching trial* was provided each group, this also taking place on trial 2, 5, 8, or 11. Finally, a control group was given 12 uninterrupted PA trials. Subjects were assigned to each group randomly, except that one subject was assigned to each group before there were any replications.

—Procedure—A 12-trial study-test procedure was used in which each study trial consisted of the presentation of the nine paired associates with a slide projector exposing each pair for 2 sec. A 60-sec test trial followed each study trial. The conventional PA tests consisted of listing the nine stimulus items and instructing the subject to write down an appropriate CVC beside each stimulus term. Four different orders of stimuli were provided. On the designated trials, either a free-recall test or a matching test replaced the conventional PA test. For the free-recall tests, the subject was instructed to write down only those CVCs he could recall; stimulus terms were not provided. For the matching tests, the response CVCs were listed in alphabetical order at the top of the sheet and the subject was instructed to write, beside each of the listed stimulus items, the CVC that had been associated with that item.

Results

The learning of the high and low intralist-similarity control groups, which were provided 12 PA trials uninterrupted by either the presentation of a single free-recall or matching test, is indicated in Figure 1 and clearly reveals the significant overall effect of intralist response similarity in paired-associate learning. Using total number of correct responses obtained over all 12 trials, a *t* of 3.52 was obtained ($p < .01$).

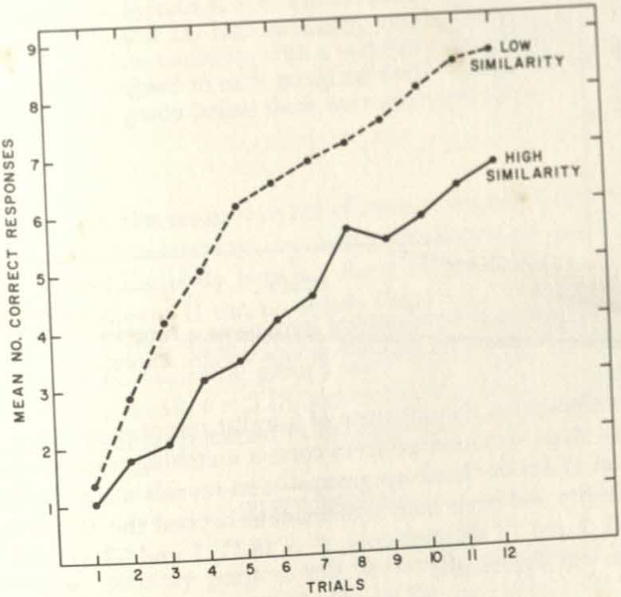


Fig. 1. Paired-associate learning as a function of high and low intralist response similarity; Experiment I

It is the influence of intralist response similarity on each of the two stages of learning, however, that is of primary interest; Figure 2 provides this information. An analysis of variance was performed using number of responses correct for the groups given a free-recall test at trial, 2, 5, 8, or 11; this was a 2 × 4 factorial with similarity and trials the two variables. The *F*s for trials (18.77, 3 and 72 *df*) and similarity (11.29, 1 and 72 *df*) were both significant beyond the .01 level; the interaction, however, was not.

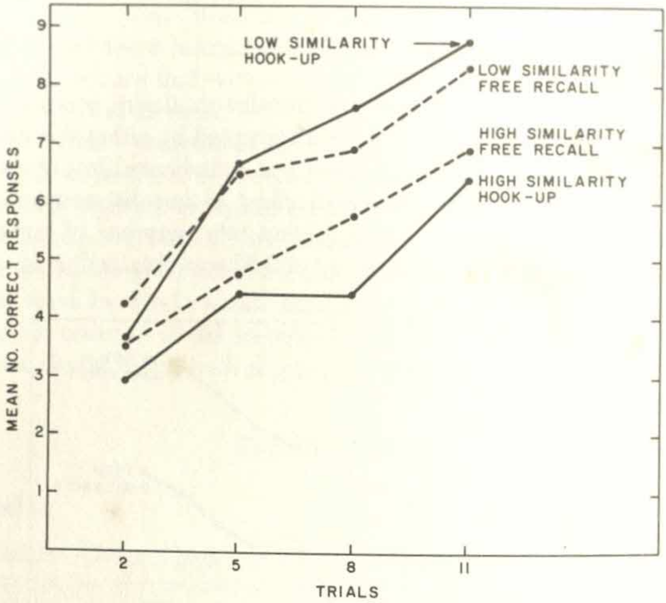


Fig. 2. Free-recall and matching-test learning as a function of high and low intralist response similarity; separate-groups design, Experiment I

An examination of the influence of intralist response similarity on the associative stage—as inferred from correct matching responses on trials 2, 5, 8, and 11 for the hook-up groups—also reveals a significant effect, with similarity and trials both significant at beyond the .01 level (trials, $F = 12.35$, 3 and 72 *df*; similarity, $F = 18.11$, 1 and 72 *df*). The interaction was not significant.

EXPERIMENT II

The experimental design used in Experiment I is costly in terms of the number of groups required to examine even a reasonable number of points along the trial continuum. Experiment II was conducted in order to determine if it would be possible to provide the experimental groups with more than a single free-recall or matching test without changing the general nature of the relationship obtained in Experiment I. More specifically, groups given high- or low-similarity material would be provided *four* free-recall or matching tests, the location of these being identical to those used in Experiment I, namely, on trials 2, 5, 8, and 11.

Method

—Materials and procedure—The same materials and procedures that were used in Experiment I were employed in Experiment II. The basic difference, of course, was that instead of utilizing four recall and four associative-learning groups for each list, just two groups were used with each list.

—Subjects—Four groups of subjects, 10 per group, were obtained from an introductory psychology course. These groups were as follows: high similarity, with a free-recall test at trials 2, 5, 8, and 11; low similarity with a free-recall test at trials 2, 5, 8, and 11; high similarity, with a matching test on trials 2, 5, 8, and 11; and low similarity, with a matching test on trials 2, 5, 8, and 11. Subjects were assigned to each group randomly except that one subject was assigned to each group before there were any replications.

Results

Figure 3 reveals the mean number of correct responses for the free-recall and associative-learning (hook-up) conditions on trials 2, 5, 8, and 11. A marked similarity between the findings obtained in Experiment I and Experiment II can be noted. The t tests between the high- and low-similarity free-recall groups and between the high- and low-similarity associative-learning groups were both significant at beyond the .01 level (free recall, $t = 3.13$; associative learning, $t = 2.89$), thus confirming the findings obtained in Experiment I.

DISCUSSION

Although our primary purpose was to examine the role of high and low formal intralist response similarity on the two stages that Underwood and Schulz (1960) have hypothesized to characterize paired-associate learning, our findings strongly support the more general position that the learning is a function of the amount of formal similarity present in the list. In all comparisons of appropriate groups provided high and low similarity, the group given low-similarity items to learn did so significantly more rapidly than the group given high-similarity responses.

Our basic concern, however, was to examine how intralist response similarity influences the two stages of paired-associate learning. First,

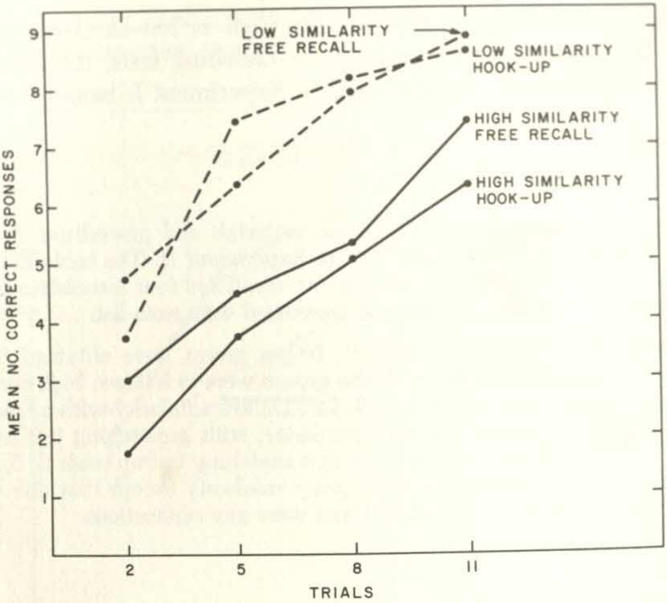


Fig. 3. Free-recall and matching-test learning as a function of high and low intralist response similarity; each group given four matching-test or free-recall trials, Experiment II

our results indicate that response learning is facilitated by low similarity, with the superiority of this group being manifested throughout the 12 trials. Underwood, Ekstrand, and Keppel (1964) and Stimmel and Stimmel (1967, 1968) have obtained similar findings.

It should be noted that this superior response learning with low formal similarity of the response items is in contrast to the results reported by both Jung (1965) and Horowitz (1961). Jung has reported that such similarity is not a significant variable in the response-learning stage, while Horowitz (1961) obtained a free-recall measure with a serial-learning task and has reported that *high* formal similarity facilitates learning. As Underwood, Ekstrand, and Keppel (1964) have pointed out, however, it was possible for Horowitz's subjects to discover rules by which many of his highly similar CVCs could be generated. Finally, Underwood, Runquist, and Schulz (1959) have also reported that response learning occurs more rapidly for highly similar items than for low. The facilitative effect reported by these authors, however, is not surprising, since they were dealing with meaningful or semantic rather than formal similarity. There undoubtedly existed high interitem associative strength among the high-similarity items (synonyms) they used, and Deese (1959) has shown that this variable greatly facilitates free recall. The facilitative effect produced by the low-similarity items in the present study undoubtedly arises, as Underwood (1966) has suggested, from the fact that the recall of low-meaningfulness CVCs requires an associative phase; that is, the subjects have to learn the associative connections between the letters that make up each item. The learning of these connections is inhibited to a much greater extent by the highly similar items.

If we can infer associative learning from the findings obtained with the matching test, our results parallel those obtained with response learning and indicate that associative learning is also a function of intralist response similarity. Horowitz (1962) has reported similar findings. The difficulty of discriminating among the high-similarity response items undoubtedly contributes to the poorer matching-test scores.

A comparison of the amount of response and associative learning has led Jung (1965) to report that response learning is greater than associative learning at the beginning of the task. On the other hand, Horowitz and Larsen (1963) have reported that the associative-learning stage takes place more rapidly than the response-learning stage. Our findings support neither position. On trial 2, the first trial in which such a comparison could be made, free-recall scores were not significantly different from matching scores, either in Experiment I or II.

Considerable question must be raised, however, as to whether or not this kind of comparison between response measures is appropriate. The situation is not unlike the traditional comparison between recall- and recognition-test scores. It has been demonstrated by various investigators that any difference between these measures is a function of the conditions under which the measurement is obtained; hence, any general conclusion about the rapidity with which response learning and associative learning take place would appear to be open to question.

The last point we should like to make is that the use of a single group provided multiple free-recall or matching tests appears to provide experimental findings similar to those obtained when multiple groups are employed. Although this finding should be verified in a variety of other experimental conditions, it could, if confirmed, effect considerable savings in the number of subjects necessary to conduct this type of experiment. One precaution should probably be exercised with this type of design: sufficient numbers of PA tests should be provided in order to ensure that the subjects perceive their primary task as one of learning the paired associates.

Notes

Received for publication October 5, 1970.

References

Deese, J. 1959. Influence of inter-item associative strength upon immediate free recall. *Psychological Reports* 5:305–312.

Goss, A. E., and Nodine, C. F. 1965. *Paired-associates learning: The role of meaningfulness, similarity, and familiarization*. New York: Academic Press.

Horowitz, L. M. 1961. Free recall and ordering of trigrams. *Journal of Experimental Psychology* 62:51–57.

Horowitz, L. M. 1962. Associative matching and intralist similarity. *Psychological Reports* 10:751–757.

Horowitz, L. M., and Larsen, S. R. 1963. Response interference in paired-associate learning. *Journal of Experimental Psychology* 65:225–232.

Jung, J. 1965. Two stages of paired-associate learning as a function of intralist similarity (IRS) and response meaningfulness (M). *Journal of Experimental Psychology* 70:371–378.

Noble, C. E. 1961. Measurements of association value (*a*), rated associations (*a'*), scaled meaningfulness (*m'*) for the 2100 CVC combinations of the English alphabet. *Psychological Reports* 8:487–521.

Stimmel, D. T., and Stimmel, N. S. 1967. Free recall as a function of intralist similarity and order of presentation constraints. *Psychological Reports* 21:541–544.

Stimmel, D. T., and Stimmel, N. S. 1968. Free recall of low meaningfulness trigrams as a function of four levels of intralist similarity. *Psychological Reports* 23:895–898.

Underwood, B. J. 1953. Studies of distributed practice: VIII. Learning and retention of paired nonsense syllables as a function of intralist similarity. *Journal of Experimental Psychology* 45:133–142.

Underwood, B. J. 1966. *Experimental psychology*. 2d ed. New York: Appleton-Century-Crofts.

Underwood, B. J., Ekstrand, B. R., and Keppel, G. 1964. Studies of distributed practice: XXIII. Variations in response-term interference. *Journal of Experimental Psychology* 68:201–212.

Underwood, B. J., Runquist, W. N., and Schulz, R. W. 1959. Response learning in paired-associate lists as a function of intralist similarity. *Journal of Experimental Psychology* 68:70–78.

Underwood, B. J., and Schulz, R. W. 1960. *Meaningfulness and verbal learning*. Philadelphia: Lippincott.

American Journal of Psychology
1971, Vol. 84, No. 4

MODALITY, DIFFICULTY, AND 'COUPLING' IN VIGILANCE BEHAVIOR

Michel Loeb and John R. Binford
University of Louisville

With auditory, loosely coupled visual, and closely coupled visual intensity-discrimination tasks matched for difficulty (*d'*) for each of 20 subjects in a short-term monitoring session, subsequent vigilance performance on the tasks was predictable, and a number of indices of performance and performance change correlated significantly and appreciably across the vigilance tasks. There were indications of modality specificity, but degree of coupling was not as significant a factor as expected.

In general, correlations of performance on different vigilance tasks have been reported to be quite low (Buckner and McGrath, 1963), especially when the tasks have involved different sensory modalities. This finding poses problems for most of the conventional theories of monitoring behavior, which do not identify task-specific or modality-specific factors.

One explanation of the low correlations may be, as Hatfield and Loeb (1968) pointed out, that most studies comparing vigilance tasks in different modalities have confounded modality with task difficulty and degree of coupling. On the matter of task difficulty, there have in some cases (not all) been efforts to equate the task in terms of some index of performance—generally, percentage of detected signals. Still, if subjects adopt different criteria for responding on different kinds of tasks, equating on this basis may not be reasonable (Green and Swets, 1966).

On the matter of coupling, a definition is in order. 'Coupling,' a term first employed by Elliott (1960), refers to the degree to which subjects are 'tied in' to the display on which the signal is presented, irrespective of their orienting or observing behavior. Generally, subjects are said to be closely coupled to an auditory display, especially one presented through earphones, and loosely coupled to a visual display, especially one requiring foveal acuity. It is also possible to devise a closely coupled visual display by presenting signals through changes in room illumination.

The results have not been conclusive. To put them in a common notation, the short-term monitoring experiment of Gunn and Loeb (1967; an attempt to match in average sensitivity an auditory and a closely coupled visual task) found moderately high correlations for the theory of signal detection indices *d′* and β as well as for false positive (or 'false alarm,' *FA*) responses. Detections (or 'hits,' *D*) were not very highly correlated in this experiment. The long-term monitoring experiment of Hatfield and Loeb (1968; an attempt to match in average discriminability an auditory, a closely coupled visual, and a loosely coupled visual task) found significant and appreciable correlations between the three conditions for *D* and, under some conditions, for *FA*, *d′* and β. For some measures, coupling seemed to be a major factor; for others, sensory modality.

A different explanation of the reported failures of correlation has recently been advanced by Sverko (1968). He noted that most, if not all, of the experiments reporting negligible correlations involved the monitoring of near-threshold signals, and he suggested that if sensory acuity were not a factor, one might expect higher correlations. In his experiment the subjects observed (in separate sessions) auditory, visual, and electrocutaneous trains of pulses and were told to respond whenever the number of pulses in successive trains was the same. Under these circumstances, correlations between *D* and *FA* in the different modalities were rather high (ranging from .66 to .84, only slightly lower than the task reliabilities).

It could be argued that in Sverko's task the subjects were reacting to numerosity, a characteristic essentially independent of the modality in a way that other, more primary sensory qualities—say, intensity and pitch or color—cannot be. Still, this is simply hypothesis; intensity might equally well be a quality transcending modality.

The present experiment was designed to determine whether the correlations for visual and auditory intensity-discrimination tasks would be higher if task difficulty were *individually* matched, for each subject, and to determine whether the Hatfield and Loeb results would be replicated under these conditions.

METHOD

—Subjects—The subjects were 20 students at the University of Louisville, paid for their participation in the experiment. Each underwent one selection and training session, three matching sessions, and three vigilance sessions.

—Selection and training session—In the first session the subjects' auditory thresholds were measured with a clinical audiometer, and those with thresholds more than 20 db above 0 (ASA) were rejected. The subjects were then trained for 10 min each on an auditory, a closely coupled visual, and a loosely coupled visual monitoring task.

The training on the *auditory task* consisted, first, of asking each subject to listen to a train of white noise, .5-sec pulses presented through earphones. The interval between initiation of successive pulses was 2.5 sec. Threshold for the pulse train was determined by a method-of-limits procedure and intensity of the pulses was set 60 db above threshold. The subject was then told that his task was to listen to the pulse train and to detect signals—noise pulses, infrequently presented, slightly more intense than the usual pulses—and respond as rapidly as possible with a key press. In this part of the training session, signal level was 61.7 db, 1.7 db greater than a nonsignal pulse. Signal rate was 2 per min.

During training on the *closely coupled visual task*, each subject sat in a slightly reclining chair with his head relatively fixed in a headrest. His eyes were sealed shut with transparent tape, and he was asked to observe a series of successive room illuminations produced by periodic lighting (by fluorescent tubes at the back of a light box) of a translucent Plexiglass square, 24 by 24 in. at a distance of 38 in. Intensity at his head was 3 ftc. Timing of the pulses was the same as in the auditory task. He was to respond with a key press to signals 6.3 ftc in intensity.

Training on the *loosely coupled visual task* was the same as on the closely coupled one, except that the subject's eyes were open and the observed source of illumination was a translucent Plexiglass disc 2 in. in diameter. Base-level pulses were 40, higher-level pulses 84 ftL. Whereas in the closely coupled task the room was dark between pulses, in this loosely coupled task the room illumination was 15 ftc.

In all these training tasks an average of two signals were presented each minute; interstimulus intervals ranged from 10 to 50 sec. Each subject performed 10 min with verbal feedback and 10 min without, on each of the three tasks. A 5-min rest was given between conditions.

—Matching sessions—In the *auditory* matching session, the levels were 1.4, 1.7, or 2.0 db above the levels of the nonsignal pulses. In the *closely coupled visual* task, the signal pulses were 4.5, 6.3, and 7.5 ftc (at the subject's head position), that is, at levels 1.8, 3.0, and 4.0 db above the nonsignal pulses. In the *loosely coupled visual* task, the intensities of the spot when signals were present were 60, 84, and 100 ftL—db increments identical to those in the closely coupled visual tasks.

The measures *D* and *FA* were defined respectively as the correct responses within 2.0 sec after signal discrimination (before the next nonsignal pulse) and as the responses at other times. These data were tabulated and from them the signal-detection indices d' and $\log \beta$ were plotted. The intensity levels with the most closely matched d' indices for the three tasks were selected for use in the later experimental (vigilance) sessions. In case of uncertainty as to which two levels were most closely matched, $\log \beta$ values were used to resolve the uncertainty.

—Vigilance sessions—The final sessions were the vigilance sessions, during which each subject was run on three separate days, one task per day, at the most similar d' levels established in the matching sessions. Within each 1-hour session there were 120 signals, 30 within each 15-min block. Before each session there was a brief warm-up period in which each subject was allowed to detect three signals. The measures for D and FA were scored as in matching sessions.

RESULTS

Of interest are the relationships of performance on the different tasks in the short-term (matching) sessions, the relationship of performance in each short-term session to that in the long-term (vigilance) session for the same task, and the intercorrelations of performance measures in the vigilance sessions.

Matching sessions

Table 1 shows the matching-session correlations between the different but difficulty-matched tasks for each of the vigilance measures—D, FA, d', and log β. Note that these correlations are generally moderately high and usually highly significant ($p < .01$).

Table 1. Correlations of performance (D, FA, d', and log β) between difficulty-matched auditory (A), closely coupled visual (VC), and loosely coupled visual (VL) tasks; matching sessions

	D		FA		d'		Log β	
	VC	VL	VC	VL	VC	VL	VC	VL
A	$.55^{**}$	$.59^{**}$	$.66^{**}$	$.61^{**}$	$.88^{**}$	$.85^{**}$	$.46^{*}$	$.55^{*}$
VC		$.59^{**}$		$.66^{**}$		$.80^{**}$		$.59^{**}$

$^{*}p < .05$. $^{**}p < .01$.

Analyses of variance for these measures in the matching sessions indicate that D varied significantly across different conditions ($p < .025$), detections being significantly fewer in the auditory condition. However, d' (the parameter on which levels of difficulty were matched) did not vary significantly across conditions; nor did FA. Log β was a significant source of variance, being highest in the auditory condition and lowest in the loosely coupled visual condition. These relationships are shown in the bar graphs at the right of Figures 1–4.

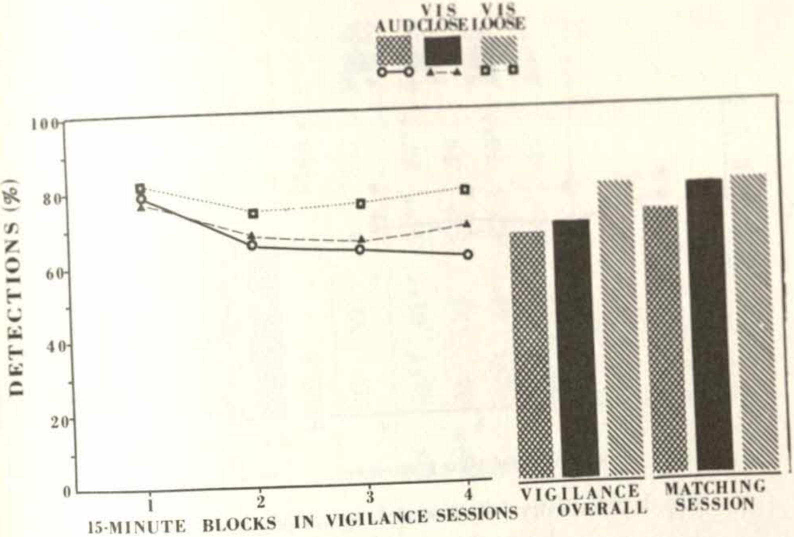


Fig. 1. Detections (*D*) in matching and vigilance sessions

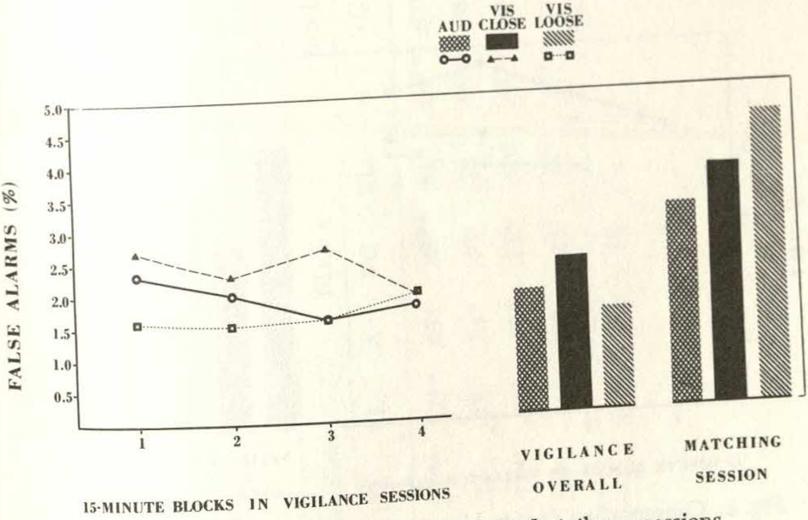


Fig. 2. False alarms (*FA*) in matching and vigilance sessions

Matching versus vigilance

Consider next the relationship of performance in the matching session to that in the vigilance session. Table 2 gives the correlation between

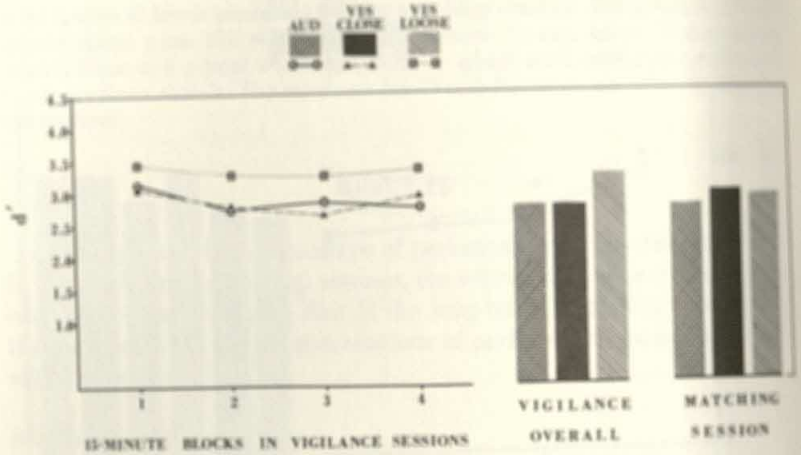


Fig. 3. Sensitivity (d') in matching and vigilance sessions

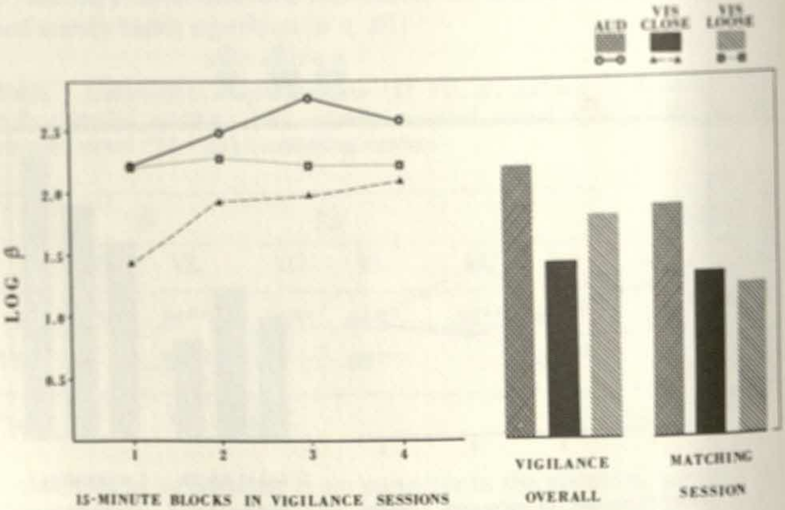


Fig. 4. Conservatism (natural log β) in matching and vigilance sessions

matching-session and vigilance-session measures, both for the entire vigilance session and for each 15-min block within it.

For the entire vigilance session, D was significantly correlated (with matching-session D) for every task, being highest in the closely coupled

Table 2. Correlations of performance (*D*, *FA*, *d'*, and log β) on auditory (A), closely coupled visual (VC), and loosely coupled visual (VL) tasks between the matching session and the vigilance session; entire vigilance session and each 15-min block thereof

	Entire session			Block 1			Block 2			Block 3			Block 4		
	A	VC	VL	A	VC	VL	A	VC	VL	A	VC	VL	A	VC	VL
D	63**	.84**	.57**	.45*	.88**	.51*	.64**	.63**	.41+	.54*	.83**	.61**	.56**	.67**	.50**
FA	57**	.30	.23	.54*	.43+	.38+	.56**	.10	.16	.56**	.35	.13	.51*	.29	.22
d'	61**	.71**	.52*	.50*	.71*	.50*	.67**	.55*	.37	.46*	.69**	.49*	.40+	.70**	.60**
Log β	50*	.39+	.48*	.33	.45*	.27	.46*	.36	.39+	.36	.44*	.47*	.42+	.17	.59**

*$p < .05$. **$p < .01$. +$p < .10$.

and lowest in the loosely coupled visual task. For *FA* a significant correlation was obtained only in the auditory task. For d' all correlations were significant; for log β correlations were moderate and significant for the auditory task and the loosely coupled visual one; for the closely coupled visual task the correlation was moderate and approaching significance.

For the first block of the vigilance session, the correlations followed a similar pattern. For log β, however, only the correlation for the closely coupled visual task was significant. On the blocks following, there may have been a slight tendency for the correlations generally to become successively lower, but there were numerous exceptions and we cannot state that this occurs in any systematic fashion.

A comparison was made of the mean performance measures in the matching session and the first vigilance block. (Analyses of variance indicated no significant difference between the matching session and the *entire* vigilance session on the comparisons involving percent *D*.) The *FA* percentages declined significantly between the matching session and the first vigilance block ($p < .01$); a significant interaction of sessions (matching versus vigilance) and conditions was noted ($p < .05$), the decline being greatest in the loosely coupled visual condition. On the other hand, d' was significantly greater in the first vigilance block than in the matching session, as was log β ($p < .05$ for both). For the latter measure a significant difference between conditions was also noted, log β being largest for the auditory task and smallest for the closely coupled visual task. Generally the analyses further indicated that percent *D* was significantly less for the first vigilance block than for the matching session ($p < .01$), especially for the loosely coupled visual condition (interaction of matching versus first-block vigilance session and conditions significant at the .01 level).

Vigilance sessions

Correlations of the different performance measures for the entire vigilance session and each block for each task are shown in Table 3. (Scatterplots for the correlations were examined. There appeared to be little or no evidence of curvilinearity in the relationship—little or no probability that the correlations were due to extreme values.)

Note that for *D* the correlations were generally sizable and significant. For *FA* there appeared to be generally significant and appreciable intertask correlations throughout the session only for the two visual tasks (closely and loosely coupled); other correlations tended to be significant only in the first block. For d' there tended to be significant correlations

Table 3. Intercorrelations of performance (D, FA, d', and $\log \beta$) between auditory (A), closely coupled visual (VC), and loosely coupled visual (VL) tasks; entire vigilance session and each 15-min block thereof

		Entire session		Block 1		Block 2		Block 3		Block 4	
		VC	VL	VC	VL	VC	VL	VC	VL	VC	VL
D	A	$.65^{**}$	$.67^{**}$	$.86^{**}$	$.51^{*}$	$.67^{**}$	$.47^{*}$	.18	$.51^{*}$	$.46^{*}$	$.58^{**}$
	VC		$.69^{**}$		$.70^{**}$		.28		$.74^{**}$		$.67^{**}$
FA	A	.15	.29	$.45^{*}$	$.60^{**}$	.09	.17	.02	.18	.11	.13
	VC		$.91^{**}$		$.75^{**}$		$.90^{**}$		$.83^{**}$		$.84^{**}$
d'	A	.36	$.38^{+}$	$.59^{**}$	$.43^{+}$	$.58^{**}$	.16	.01	.25	.14	.22
	VC		$.56^{**}$		$.54^{*}$		.21		.35		$.70^{**}$
$\log \beta$	A	$.56^{**}$	$.72^{**}$	$.54^{*}$	$.45^{*}$	$.52^{*}$	$.70^{**}$	.19	$.52^{*}$	$.45^{*}$	$.69^{*}$
	VC		$.73^{**}$		$.70^{**}$		.53		$.70^{**}$		$.42^{+}$

$^{*}p < .05$. $^{**}p < .01$. $^{+}p < .10$.

overall and throughout the sessions. This is somewhat true for log β, and there appears to have been some tendency for correlations to decrease over time.

Analyses of variance were also performed for each measure within vigilance sessions. Generally, *D* declined significantly, as in all vigilance sessions ($p < .01$). There was significant intertask variance, *D* being highest in the closely coupled visual condition. *FA* did not decline significantly. This is unusual in vigilance experiments; the probable explanation is that the learning effect usually occurring within sessions instead took place in the matching sessions. The analyses indicated a significant decrement in d' ($p < .025$) and significant intertask variance ($p < .025$), with d' greatest in the loosely coupled visual condition. Log β increased significantly ($p < .025$) and showed significant intertask variance ($p < .01$), with log β greatest in the closely coupled visual condition and least in the auditory condition.

Another question concerns the correlations of trends between conditions. Will a subject who has a large decrement on one vigilance task have a similar decrement on another? To answer it, percent changes (amount of change between blocks 1 and 4, divided by the value in block 1, with direction of change taken into account) were computed for each dependent measure and each subject, and correlations for these were obtained. Note, in Table 4, that moderately large, statistically significant correlations for *D* were obtained, correlations ranging from .48 to .71. On other measures only one correlation was significant (on *FA* for closely versus loosely coupled visual tasks, $r = -.50$, $p < .05$), and this was not in the expected direction.

Table 4. Intercorrelations of percent change in performance (*D*, *FA*, d', and log β between auditory (A), closely coupled visual (VC), and loosely coupled visual (VL) tasks; vigilance session, blocks 1 and 4

	D		*FA*		d'		log β	
	VC	VL	VC	VL	VC	VL	VC	VL
A	.48*	.55**	−.28	.20	.17	.10	−.26	.38
VC		.71**		−.50*		.16		.29

*$p < .05$. **$p < .01$.

DISCUSSION

The moderately high correlations obtained in the matching session (Table 1) simply demonstrate that if subjects are fairly well matched in terms of sensitivity (d'), they are reasonably well matched in terms of primary performance measures (D and FA) as well as the other signal-detection measure (β).

It was anticipated that if the matching-session scores (for specific types of tasks) were correlated with those in the vigilance sessions, then the highest correlations would be between the matching-session measures and those in the first 15 min of the vigilance session (when the subjects presumably would be maximally alerted). No such trend emerged (Table 2). Generally moderate correlations prevailed throughout, though for the FA measure in the closely and loosely coupled visual conditions the correlations were generally low and nonsignificant.

The intercorrelations of performance on the three tasks in the vigilance session (Table 3) were moderately high, especially for the D and log β measures. In block 1 the intercorrelations were generally rather high for all measures; but for the entire session and for the remaining blocks, this was not generally true for the FA and d' measures. For the entire session and for blocks 2, 3, and 4, FA correlated well only for the two visual conditions. For the entire session and for block 4, d' correlated significantly only for the visual conditions; in block 3 no significant correlations were obtained for d' at all, and in block 2 the significant correlation was for the closely coupled tasks (for the auditory and the closely coupled visual tasks) rather than for both the visual ones. These data, unlike those in Table 2, suggest that there are factors specific to alerted conditions, as well as some specific to modality and, perhaps, coupling.

The correlations of percent change (Table 4) indicate that changes in D only are predictable from one task to another. It may be of some significance that the largest correlation was that between the two visual conditions. (A negative correlation of .50, between percent change of FA in the closely and loosely coupled visual conditions was obtained. However, the correlations of FA changes between the other conditions were dissimilar, and since there seems little reason to expect such a correlation, this probably should be interpreted as a chance finding.) Since the tasks were initially matched in terms of d' rather than D, it is a little surprising that only the changes in D were significantly correlated. This may simply reflect the fact that, on the average, the changes in D were greater than in other measures, for although a number of studies have shown changes in D and FA, changes in d' have not always occurred and or have usually been small. The best conclusion from the correlations in Table 4 is that only D seems to exhibit a correlation of changes within sessions.

The data do not entirely support Sverko's suggestion (1968) that correlations of performance can be obtained only for suprathreshold tasks. They indicate that when performance in the different modalities is equated under alerted conditions, then there will be high correlations on certain indices of overall performance and changes in performance. However, Sverko felt that the low correlations usually obtained may reflect differing sensory acuities in different modalities, and the procedure employed in this experiment largely eliminates such differences.

In conclusion, if vigilance tasks in auditory and visual modalities are equated for individual subjects in terms of type and difficulty of discrimination, then

1. Performance measures [detections, or hits (D); false alarms, or false positives (FA); sensitivity (d'); and criteria ($\log \beta$)] are rather highly correlated for a short (15-min) monitoring session;
2. D, d', and $\log \beta$ in a longer (1-hour) session can be predicted rather well from performance in the short session. For FA this appears to be true only for the auditory task;
3. D and $\log \beta$ are rather highly correlated across tasks for a 1-hour vigilance session. FA and d' correlate significantly only for two visual conditions, one closely coupled and the other loosely coupled, though there appear to be lower correlations between the other conditions, especially for d'. In the first 15 min of the vigilance session almost all measures are correlated across tasks; and
4. Changes in D, generally decrements within sessions, are appreciably correlated across tasks.

The results, taken as a whole, suggest that when tasks in different modalities are matched for difficulty, then performance (and changes in performance) on them will be correlated—that monitoring behavior is not entirely modality specific. On the other hand, the data also suggest that there probably are modality-specific elements involved too, as performance on the visual tasks tended to correlate most highly. There is little evidence that degree of coupling played a role in the present experiment; still, using an even more loosely coupled visual task (say, one requiring a very precise spatial discrimination or constant foveal viewing) might alter that outcome.

Notes

The authors are indebted to Lois H. Benedetti, Ernest M. Weiler, and Carlyn Fleck for their technical assistance. This research was supported by a contract between the Office of the Surgeon General, U. S. Army Medical Research and Development Command, and the University of Louisville. Requests for reprints should be mailed to the first author at the University of Louisville, Department of Psychology, Louisville, Kentucky 40208. Received for publication December 21, 1970.

References

Buckner, D. N., and McGrath, J. J. 1963. A comparison of performance on single and dual sensory mode vigilance tasks. In *Vigilance: A symposium*, ed. D. N. Buckner and J. J. McGrath, pp. 53–71. New York: McGraw-Hill.

Elliott, E. 1960. Perception and alertness. *Ergonomics* 3(4):357–364.

Green, David M., and Swets, John A. 1966. *Signal detection theory and psychophysics*. New York: Wiley.

Gunn, Walter J., and Loeb, Michel. 1967. Correlation of performance in detecting visual and auditory signals. *American Journal of Psychology* 80:236–242.

Hatfield, Jimmy L., and Loeb, Michel. 1968. Sense mode and coupling in a vigilance task. *Perception and Psychophysics* 4(1):29–36.

Sverko, B. 1968. Intermodal correlations in vigilance performance. *Proceedings of the XVIth International Congress of Applied Psychology, Amsterdam.*

continued from page 500

Reagles, Kenneth W.; Wright, George N.; and Butler, Alfred J. *Correlates of Client Satisfaction in an Expanded Vocational Rehabilitation Program.* Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1970. Pp. xiii, 57.

Riener, Scholomo. *The Quest for Security: A Voyage of Exploration into the Vast Reaches of Inner Space.* Jerusalem, Israel: The Psychonauts, 1971.

Robertson, James. *Young Children in Hospitals.* 2d ed. New York: Barnes & Noble, 1971. Pp. xv, 155. $5.25.

Rosenfels, Paul. *Homosexuality: The Psychology of the Creatice Process.* Roslyn Heights, N.Y.: Libra Publishers, 1971. Pp. 169. $5.95.

Ross, Jack C., and Wheeler, Raymond H. *Black Belonging: A Study of the Social Correlates of Work Relations among Negroes.* Westport, Conn.: Greenwood Publishing, 1971. Pp. xii, 292.

Rothman, G. *The Riddle of Cruelty.* New York: Philosophical Library, 1971. Pp. 210. $7.50.

Rotter, Julian B. *Clinical Psychology.* 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. x, 117.

Sacks, Oliver W. *Migraine: Evolution of a Common Disorder.* Berkeley and Los Angeles: University of California Press, 1971. Pp. 298. $8.50.

Schroder, Harold M., and Suedfeld, Peter, eds. *Personality Theory and Information Processing.* New York: Ronald Press, 1971. Pp. v, 303. $8.50.

Smith, C. U. M. *The Brain: Towards an Understanding.* New York: G. P. Putnam's Sons, 1970. Pp. 392. $10.00.

Steger, Joseph A., ed. *Readings in Statistics for the Behavioral Scientist.* New York: Holt, Rinehart and Winston, 1971. Pp. ix, 406.

Stutz, Robert M.; Dember, William N.; and Jenkins, James J., eds. *Exploring Behavior and Experience: Readings in General Psychology.* Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xiii, 366. Paperback, $5.95.

Sutherland, N. S., and Mackintosh, N. J. *Mechanisms of Animal Discrimination Learning.* New York: Academic Press, 1971. Pp. xi, 559.

Triandis, Harry C. *Attitude and Attitude Change.* New York: Wiley, 1971. Pp. xiii, 232. $6.95.

Tulipan, Alan B., and SiSalvo, Charles, eds. *The Psychiatric Clinic in Encounter.* Oil City, Pa.: POCA Press, 1971. Pp. xiv, 194. $6.95.

Tyler, Leona E. *Tests and Measurements.* 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xii, 99. $5.75; paperback, $2.75.

Veitzman, Vicente. *Psicofarmacologia clinica aplicada.* Buenos Aires: Lopez Libreros Editores, 1970. Pp. xiv, 294.

Wittenborn, J. R.; Smith, Jean Paul; and Wittenborn, Sarah A., eds. *Communication and Drug Abuse.* Springfield, Ill.: Charles C Thomas, 1971. Pp. xiii, 542. $23.25.

Wright, George N.; Reagles, Kenneth W.; and Butler, Alfred J. *An Expanded Program of Vocational Rehabilitation Methodology and Description of Client Population.* Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1971. Pp. xi, 148.

Zusne, Leonard. *Visual Perception of Form.* New York: Academic Press, 1970. Pp. xi, 547.

American Journal of Psychology
1971, Vol. 84, No. 4

TEST OF A MULTIPLYING MODEL FOR ESTIMATED AREA OF RECTANGLES

Norman H. Anderson and David J. Weiss
University of California, San Diego

Graphic ratings of physical size were obtained for 36 rectangles in a 6 × 6, width × height design, with each factor ranging from 3 to 18 cm. These judgments were approximately linear in physical area and followed a multiplying model reasonably well, though not perfectly. It is suggested that the underlying process was one of additive integration and that functional-measurement procedures can be used to scale phenomenal size of complex shapes.

Previous work on judgments of area has been largely concerned with accuracy, and with the relation between the response and the physical area (Ekman and Junge, 1961; Stevens and Guirao, 1963; Teghtsoonian, 1963; Stanek, 1969). This orientation toward the physical stimulus measure has tended to pass over the psychological processes involved in area judgments. The importance of a process orientation is suggested by reports that judged area appears to be affected by shape as well as by physical area (Luckiesh, 1922, p. 97; Warren and Pinneau, 1955; Smith, 1969). An interesting example is cited by Paterson and Tinker (1938; see also Helson and Bevan, 1964; Tolansky, 1964, p. 49), who noted that, contrary to appearances, the marginal white space on a printed page is roughly equal to the central printed area. In a typical page of text in this journal, for example, the marginal area occupies approximately 44% of the page.

The experiment reported here was designed to test a simple multiplying model for area of rectangles. The judged area was assumed to be, in effect, the product of the subjective values of width and height. A novel feature is the use of procedures from the theory of functional measurement (Anderson, 1970), procedures which test the model while allowing for subjective values of width and height. In this theory, no a priori assumption about the relation between subjective and objective stimulus values is needed. Differential weighting of the two stimulus dimensions is also allowed for, though this cannot be explicitly tested in a simple multiplying model.

METHOD

—Subjects and apparatus—The subjects were instructed to judge the actual physical area of rectangles using a linear graphic rating. The 36 basic rectangles were constructed from a symmetrical 6×6 factorial design. Both width and height of the rectangles varied from 3 cm to 18 cm in equal steps. The subject sat 1 m from the stimulus-response display and responded by varying the extent of exposed white part on an endless motorized strip of tape. The exposed white part could vary from 0 to 50 cm, and the left and right ends were anchored by squares of 1 and 23 cm respectively. The stimulus rectangle was presented midway between the two anchor cards and all these were in the same plane as the response tape. The subjects were 23 members of the student community, who were paid for their services.

—Materials and procedure—The rectangles were outlined in 1.5-mm black tape on square white tagboard cards, 23 cm on a side. In addition to the 36 stimuli, filler rectangles of 1 by 2, 1 by 3, 2 by 2, 18 by 20, 18 by 21, 20 by 21, 21 by 21, and 22 by 22 cm were included. Following 10 practice trials, these 44 stimulus rectangles were presented in a different shuffled order three times for each subject. The experimenter presented the stimuli, recorded the response, and reset the white tape to zero after each response.

RESULTS

The mean judgments are plotted as a two-way design in the left panel of Figure 1. Stimulus width is on the horizontal axis, and one curve is plotted for each of the six heights. The multiplying model implies that these curves should form a diverging fan of straight lines, and this is true to a good first approximation.

However, there is some discrepancy from the model, visible in the slight downward curvature of the bottom curve and in the more marked upward curvature of the top curve. A different view of this discrepancy is seen in the right-hand panel of Figure 1, which plots the response against the physical area on the horizontal axis. This shows small but definite deviations from a linear relation.

Almost the identical pattern of results was obtained in an earlier, unreported experiment with the same apparatus, design, and general procedure. The pattern thus appears to be reliable, but no explanation is known. A number of factors can probably be ruled out, however, and these require a few remarks.

The most important possibility is that the multiplying model applies but that subjective and objective values of width and height are not linearly related. Figure 1 is thus not a satisfactory test of the model, since it uses the physical values of the stimuli. However, a strong test

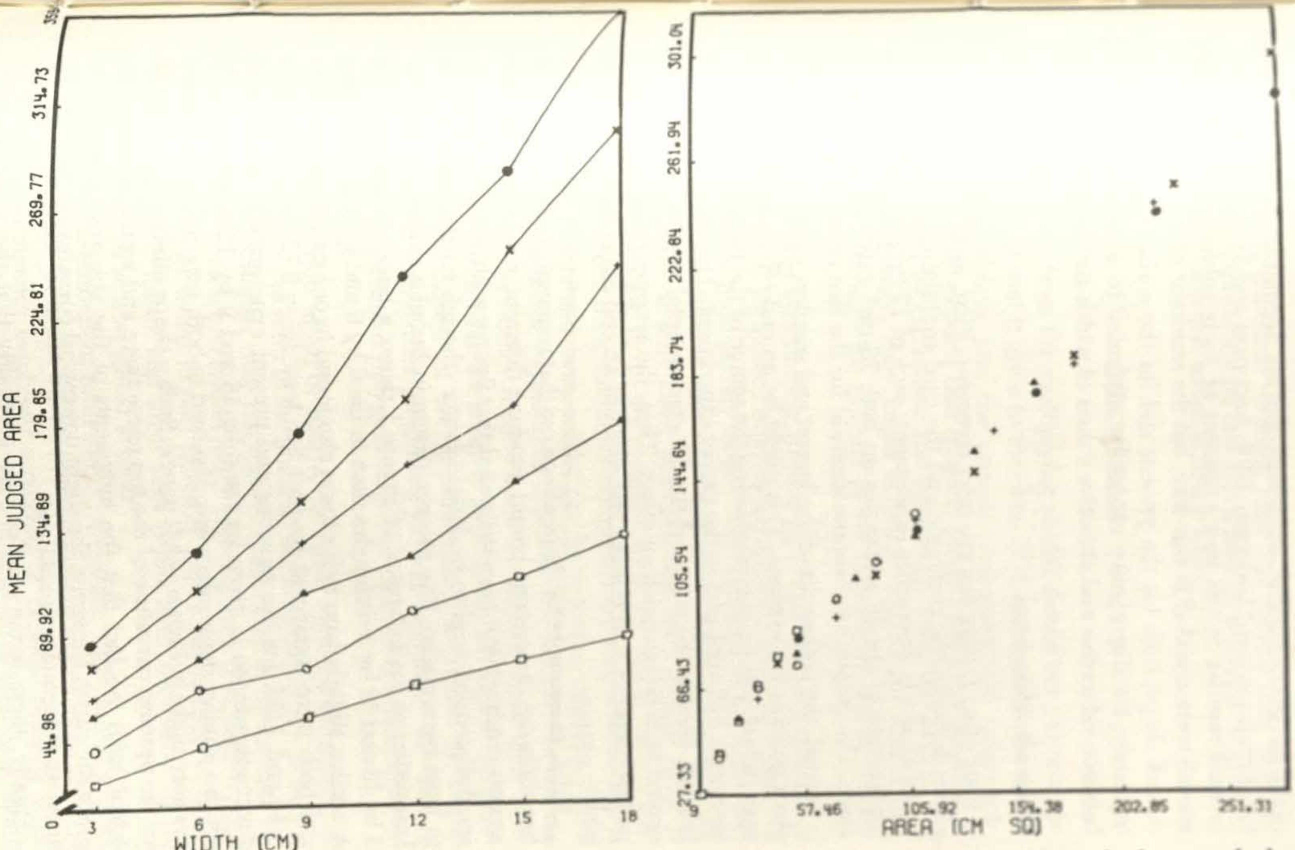


Fig. 1. Panel at the left shows mean judged area of 36 rectangles as a function of rectangle width on the horizontal; the six curves correspond to rectangles of heights 3, 6, 9, 12, 15, and 18 cm. Panel at the right shows mean judged area as a function of actual physical area on the horizontal; response was a linear graphic rating, with arbitrary zero and unit.

of the model is available, one which does not depend on the use of physical stimulus values (Anderson, 1970; Anderson and Shanteau, 1970) but instead rescales to the best estimates of their subjective values. The stimuli were rescaled in this way, but the resultant graph had the same overall shape, both for the present and for the earlier experiment. Accordingly, the discrepancy cannot be attributed to a nonlinear relation between subjective and objective values of width and height.

Other considerations

Four other factors deserve brief comment. First, assimilation-contrast effects between the rectangle and the card containing it can probably be ruled out. In the earlier experiment, each of 17 subjects judged three replications of the 36 rectangles on both 28-cm^2 cards and on 41-cm^2 cards. The graphs were almost identical for the two conditions.

Second, the judgments of the largest and smallest squares in Figure 1 seem perhaps too extreme. This might be an end effect in the response scale, which was suspected of being the source of the trouble, since similar results had appeared in the earlier experiment. The present filler rectangles, outside the range of the 36 experimental rectangles, were included to eliminate such an effect. Thus, the reappearance of the same pattern suggests that it is not the result of an end effect in the response scale.

Third, there might be shape effects, so that rectangles of the same area but different dimensions would produce different responses. Previous reports (Smith, 1969) have suggested that figures with one large and one small dimension tend to be judged larger (though no direct test seems to have been made). The present design included a direct comparison. Inspection of the left panel of Figure 1 shows a larger response to the 3 by 18 and 18 by 3 rectangles than to the 6 by 9 and 9 by 6 rectangles. A similar, slightly larger effect was also found in our earlier experiment, and both were significant by post hoc tests [$F = 6.72$ and 4.43, $df = 1/16$ and $1/22$, for the previous and present experiments]. One other direct comparison is given by the 6 by 18 and 9 by 12 rectangles, for which a similar shape effect was observed in both experiments—though it is very slight in Figure 1 (left panel). Such a shape effect could account for the upward curvature of the top curve there and would also be consistent with the fact that the judgments of the squares fit well to a quadratic curve. However, the slight downward curvature in the bottom curve would remain unexplained.

Finally, there is the ever-present possibility that the response scale is subject to a small nonlinear response bias. The present data unfortunately provide no information on this, but it may be desirable to use two somewhat different response scales in future work.

DISCUSSION

The discrepancy from the multiplying model should not obscure the reasonably good fit that it gives to the data. It is appropriate, therefore, to inquire more closely into the process underlying the judgments. Teghtsoonian (1965) reports that subjects judged the physical area of squares and circles by estimating and squaring a linear dimension. For rectangles, analogously, they could estimate width and height and multiply the two, though the subjects in our experiment did not report doing so.

An alternative, perhaps simpler, hypothesis is that the process is one of additive integration. Responses obeying an additive process would still appear to follow a multiplying model. This additive hypothesis also applies to irregular figures more directly than would a multiplying model. Accordingly, it may merit consideration even though there is some evidence against it. In its simplest form, it would imply that judgments of physical area should be independent of shape. Small shape effects, such as obtained here, might not be a serious problem, since shape variables might be expected to influence the integration to some degree. However, two other reports seem to indicate somewhat more substantial deviations from the additive hypothesis. The first is the page-margin effect mentioned in the introduction. The second is Teghtsoonian's report (1965) that judgments of both apparent size and of physical area for irregular polygons followed a power function with exponent of about .8.

A clear distinction must, of course, be made between judgments of size and area. Size can refer to the linear dimensions of a figure as well as to its area, and subjects may not always make a clear distinction between the two, even when told to judge physical area. The page-margin effect, for example, seems in part to result from a comparison of linear dimensions. Both linear and areal cues could then contribute to the response, perhaps as a weighted average.

Functional-measurement procedures (Anderson, 1970) might provide a useful approach. Stimulus figures could consist of two or more spatially separated components, combined according to factorial designs. If the adding hypothesis has any validity, then judgments of total area should be additive in the stimulus components. This could hold even though the components themselves were judged and evaluated on a different basis —and showed shape effects, for instance. The hypothesis of component addition can be tested simply and directly using the raw response measure. If the hypothesis holds, then the subjective values of the components can be estimated directly from the data on equal interval scales.[1] This would be an important result, since it would give a theoretically valid measure of the effective phenomenal size of the components. Such measures would provide a solid base for the study of the processes underlying the valuation of the components.

Notes

The work reported was supported in part by National Science Foundation Grant GB 6666, in part by a grant from the National Institutes of Mental Health to the Center for Human Information Processing, University of California, San Diego. The authors wish to thank Carlyn Joergensen for her assistance. Received for publication October 14, 1970.

[1] Functional scaling of length is illustrated in Weiss and Anderson, 1969.

References

Anderson, N. H. 1970. Functional measurement and psychophysical judgment. *Psychological Review* 77:153–170.

Anderson, N. H., and Shanteau, J. C. 1970. Information integration in risky decision making. *Journal of Experimental Psychology* 84:441–451.

Ekman, G., and Junge, K. 1961. Psychophysical relations in visual perception of length, area, and volume. *Scandinavian Journal of Psychology* 2:1–10.

Helson, H., and Bevan, W. 1964. An investigation of variables in judgments of relative area. *Journal of Experimental Psychology* 67:335–341.

Luckiesh, M. 1922. *Visual illusions*. Reprint, New York: Dover, 1965.

Paterson, D. G., and Tinker, M. A. 1938. The part-whole proportion illusion in printing. *Journal of Applied Psychology* 22:421–425.

Smith, J. P. 1969. The effects of figural shape on the perception of area. *Perception and Psychophysics* 5:49–52.

Stanek, R. J. 1969. A parametric study of surface and volume judgments. *Perception and Psychophysics* 6:16–18.

Stevens, S. S., and Guirao, Miguelina. 1963. Subjective scaling of length and area and the matching of length to loudness and brightness. *Journal of Experimental Psychology* 66:177–186.

Teghtsoonian, Martha. 1965. The judgment of size. *American Journal of Psychology* 78:392–402.

Tolansky, S. 1964. *Optical illusions*. New York: Pergamon Press.

Warren, J. M., and Pinneau, S. R. 1955. Influence of form on judgments of perceptual area. *Perceptual and Motor Skills* 5:7–10.

Weiss, D. J., and Anderson, N. H. 1969. Subjective averaging of length with serial presentation. *Journal of Experimental Psychology* 82:52–63.

American Journal of Psychology
1971, Vol. 84, No. 4

CLASSICAL RESPIRATORY CONDITIONING IN THE FISH: CS INTENSITY

W. T. Woodard
University of South Florida

Fish (*Cyprinus carpio*) were trained in a Pavlovian respiratory-conditioning situation with light as CS and shock as US. To distinguish learning from performance effects during acquisition, two groups were trained from the start with different CS intensities and then half of each group shifted to the intensity of the other. Intensity had a clearly associative effect as well as a nonassociative one.

In a variety of Pavlovian conditioning situations, performance improves with CS intensity (Gray, 1965). But does learning—as distinct from performance—improve with CS intensity? In a conditioned-suppression experiment with rats, Kamin and Schaub (1963) found that CS intensity affected acquisition in general but not (beyond the results attributable to stimulus generalization) extinction, a finding which hardly argues for much associative learning. Other failures to find an associative effect of CS intensity have been reported in extinction experiments on eyelid and GSR conditioning in man (Grant and Schneider, 1948, 1949) and on conditioned wheel-turning in rats (Kessen, 1953). It is conceivable, however, that extinction is not the best procedure to use for measuring the effect of CS intensity on learning, as the course of extinction in classical conditioning may be too rapid to provide a sufficiently sensitive measure. In the factorial experiment now to be reported, a substantial associative effect of CS intensity was found in acquisition.

METHOD

—Subjects—The subjects were 127 Japanese carp (*Cyprinus carpio*), averaging 80 mm in length. They were selected at random from groups living in large holding tanks, immediately tested, and then discarded after a single experimental session.

—Apparatus—During training, each fish was immobilized in a padded acrylic clamp that permitted unrestricted opercular (gill-cover) activity but little other bodily movement. The clamp was submerged in a water-filled tank inside a sound-attenuating chamber.

The unconditioned stimulus (US) was a 50-msec, 7-ma, d.c. shock administered through electrodes located on either side of the fish. Resting lightly against the fish's operculum was a flexible wire probe clamped in the needle holder of a crystal phonograph cartridge (Astatic 12-U). The cartridge generated an electrical signal proportional to opercular activity, and this signal was amplified and integrated over time by a Sanborn 350–3700 integrating preamplifier. The output of the integrator yielded a measure proportional to the total amount of opercular activity within any given period of time. Both the unintegrated and integrated signals were recorded on a Grass 7 polygraph. Another channel of the polygraph was used to record CS and US presentations. At all times a light diffusely illuminated the dull white interior of the experimental chamber. The two other lamps, fitted with filters of different density, provided the diffuse low- or high-intensity illumination used as the conditioned stimulus (CS). The level of illumination, measured as incident light at the chamber wall in front of the fish, was .02 ftc with no CS, .03 ftc with the low-intensity CS turned on, and .19 ftc with the high-intensity CS turned on.

—Procedure—The pattern of opercular activity normally evident in many fish is so irregular that Otis, Cerf, and Thomas (1957), who previously investigated respiratory conditioning in the fish, found it necessary to discard much of their data. But when an irregularly respiring carp is placed with its head toward the anode in a weak d.c. electric field, the variability of opercular activity is dramatically reduced.[1] This technique was used in the present experiment after preliminary work showed good opercular conditioning to be possible in fields of up to about .10 ma/cm^2. Fields of higher intensity or of opposite polarity tended to depress performance.

Accordingly, each fish was first immobilized by mild electronarcosis, achieved by placing the fish in an anodic orientation in a .35 ma/cm^2 electric field generated by passing a d.c. current between carbon electrodes lining both ends of the experimental tank. The clamp was applied, and as soon as a clamped fish was put into the testing chamber, a d.c. field of .20 ma/cm^2 quickly stabilized its respiration. Then, 8 min later, and before any trials were begun, the field was reduced to .07 ma/cm^2, the minimal value set for acceptable respiratory regulation.

Then came (*a*) a 10-min adaptation period, (*b*) 3 US-only trials with a 30-sec mean intertrial interval, (*c*) 10 CS-only (habituation) trials with a 30-sec mean intertrial interval, and (*d*) 10 CS-US trials with a 4-min mean intertrial interval. In the four experimental groups, the CS-US trials were 10 delayed-conditioning trials with a CS-US interval of 4 sec. In the four control groups, the CS-US trials were 10 pseudoconditioning trials with the CS and US presented at random during the intertrial interval, with the restriction that no CS-US contiguity of less than 90 sec be permitted.

For the first 5 CS-US trials, half the fish were trained with high CS-intensity and half with low intensity. Half the fish initially trained with high intensity were shifted after trial 5 to low intensity (group HL), and half remained at high intensity (group HH). The fish trained with low intensity were divided into corresponding LH and LL groups. The control groups were trained with the same intensities as the experimental groups, and designated HLC, HHC, LHC, and LLC respectively. During the 10 CS-only trials preceding the conditioning or pseudoconditioning trials, each animal experienced each intensity equally often and in random order.

RESULTS

The conditioned response was a suppression of opercular (respiratory) activity during the CS-US interval. All analyses were based on measures of response magnitude obtained by computing the ratio of the integrated activity in the 4-sec CS-US interval to the activity in the immediately preceding 8-sec interval. A score of .50 thus indicated no suppression and a score of .00 indicated complete suppression. For each fish, single measures of preshift and postshift performance were obtained by computing the median respiratory ratio for trials 1–5 and trials 6–10 respectively, and these measures were used for all subsequent statistical analyses. Because there were extreme scores and skewed distributions, nonparametric statistical tests were used for all analyses.

There is clear evidence of conditioning in the experimental groups. For either low- or high-intensity conditions, a preshift comparison between experimental and pseudoconditioning control groups showed significantly more opercular suppression (greater CR magnitudes) in the former (low-intensity, Mann-Whitney $U = 243.5$, $n = 38, 26$, $p < .001$; high-intensity, $U = 63$, $n = 25, 37$, $p < .001$).

A highly reliable effect of CS intensity was demonstrated by the experimental groups during the preshift period, responses of the high-intensity groups being significantly greater than those of the low-intensity groups ($U = 116$, $n = 37, 38$, $p < .001$). The difference could have been due either to greater associative strength in the high-intensity groups, or to greater response energization (a dynamistic performance effect), or to both. Since the control groups also tended to show more suppression with the high-intensity CS ($U = 206$, $n = 25, 26$, $p = .03$), an effect which could *not* have been associative, it seems reasonable to assume that the preshift results for the experimental groups may have been due at least in part to a dynamistic effect of CS intensity on performance.

Before the postshift data could be analyzed properly, it was necessary to deal with a problem arising from a preliminary analysis that showed appreciable differences between the preshift scores of experimental groups trained under like intensities (say, groups HL and HH), a result, evidently, of sampling error. Matched groups, therefore, were selected by choosing animals on the basis of their preshift performance in such a way that the distributions of median respiratory ratios for each new group (say, HL) had the same median and interquartile range as the original group of animals trained with the same CS intensity (say, HH + HL).

Data for the matched groups are plotted in Figure 1, the upper portion of which shows the experimental performance and the lower portion of which shows the control performance. The scale of the ordinate is inverted to show improvement in performance as an ascending curve (i.e., as progressively greater suppression). It should be noted that the same postshift relationships appear in a plot of the data for the original, unselected groups.

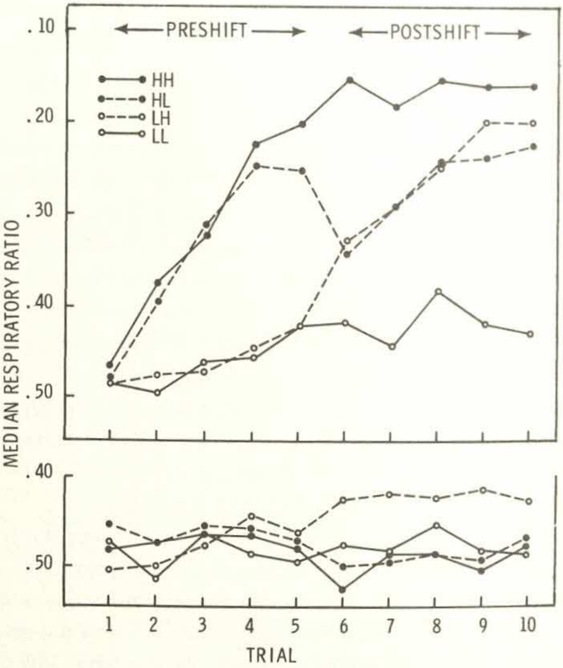


Fig. 1. Median respiratory ratios as a function of trials before and after CS-intensity shift for experimental (upper panel) and control groups (lower panel); data plotted on an inverted scale, with higher points indicating greater response to the CS

The curves for the unshifted groups show clearly that the effect of CS intensity in the early trials persisted in later trials (for LL versus HH, postshift, $U = 6, n = 12, 14, p < .002$). While the immediate postshift decrement seen in group HL (HL versus HH, Wilcoxon $T = 5, n =$

14, $p < .01$) might have been due either to reduced energization of performance or to a generalization decrement,[2] the immediate postshift increment in group LH (LH versus LL, $T = 0, n = 12, p < .01$) must have been primarily a dynamistic effect, since stimulus generalization would have tended to produce a difference in the opposite direction. There is, thus, evidence for a nonassociative effect of CS intensity.

Further evidence of a nonassociative effect is provided by the postshift control data. Although the difference between the unshifted controls was small and inconsistent, a change to a higher CS intensity clearly facilitated group LHC, whose responses were even greater than those of group HHC ($U = 4, n = 8, 8, p = .002$)—a contrast effect of the type discussed by Grice and Hunter (1964). That no such contrast effect appeared in the opposite direction was probably due to the extremely small magnitude of the responses in the low-intensity groups and the consequent measurement limitation.

Although a dynamistic effect is evident in the postshift data, a powerful associative effect of CS intensity also may be seen. Postshift responses were greater in group HL than in group LL on all trials (for postshift LL versus HL, $U = 35.5, n = 12, 14, p < .02$). This difference cannot be accounted for either in terms of response energization (since the postshift stimulus for groups HL and LL was the same) or in terms of generalization decrement (since group HL was the one for which stimulus conditions were changed), or in terms of contrast (which would predict a difference in the opposite direction for group HL). Only if the animals in group HL had formed stronger associations to the high-intensity CS before the shift could their responses to the low-intensity CS after the shift have been stronger than those of the animals in group LL, which were tested under identical conditions but trained before shift with the CS of a low intensity.

It might be desirable to offer some hypotheses explaining the differences between these results and those of the earlier factorial studies that failed to show an associative effect of CS intensity, but the present experiment was different in so many respects from the previous ones that speculation seems futile until comparative studies can be made. Finally, aside from providing evidence for an associative effect of CS intensity, the present experiment illustrates the utility of a little-used but highly efficient technique for the study of Pavlovian conditioning. The procedures described yield substantial levels of conditioning in a short time while still permitting detailed study of the course of acquisition. The technique has been shown to be quite sensitive to variation in CS intensity, a variable usually considered rather weak, and informal observations suggest that the method is equally sensitive to the effects of US intensity, CS-US interval, and intertrial interval.

Notes

This paper is adapted from a doctoral dissertation submitted to the Department of Psychology at the University of Hawaii in 1966. The writer is indebted to Professor W. F. Oakes, under whose direction the work was done.

Received for publication May 21, 1970.

1. Unpublished observations.
2. See Yarczower and Bitterman, 1965.

References

Grant, D. A., and Schneider, D. E. 1948. Intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *Journal of Experimental Psychology* 38:690–696.

Grant, D. A., and Schneider, D. E. 1949. Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *Journal of Experimental Psychology* 39:34–40.

Gray, J. A. 1965. Stimulus intensity dynamism. *Psychological Bulletin* 63: 180–196.

Grice, G. R., and Hunter, J. R. 1964. Stimulus intensity effects depend upon the type of experimental design. *Psychological Review* 71:247–256.

Kamin, L. J., and Schaub, R. E. 1963. Effects of conditioned stimulus intensity on the conditioned emotional response. *Journal of Comparative and Physiological Psychology* 56:502–507.

Kessen, W. 1953. Response strength and conditioned stimulus intensity. *Journal of Experimental Psychology* 45:82–86.

Otis, L. S., Cerf, J. A., and Thomas, G. J. 1957. Conditioned inhibition of respiration and heart rate in the goldfish. *Science* 126:263–264.

Yarczower, M., and Bitterman, M. E. 1965. Stimulus generalization in the goldfish. In *Stimulus generalization*, ed. D. I. Mostofsky. Stanford, Calif.: Stanford University Press.

American Journal of Psychology
1971, Vol. 84, No. 4

CS COMPLEXITY AND INTERTRIAL RESPONDING IN SHUTTLEBOX AVOIDANCE CONDITIONING

Donald J. Levis
University of Iowa

As CS complexity (two levels serial, two nonserial) increased, the intertrial responding of Sprague-Dawley rats decreased. 'Learner' rats did much more intertrial responding than did 'nonlearners,' and their responding correlated positively with—but had no direct effect on the absolute level of—their conditioned avoidance responding, increasing over trials and then tapering off. The general findings, replicated with Iowa hooded rats, imply that intertrial responses are generalized avoidance responses contingent upon CS-US pairing and subject to CS discriminability.

Responses during the intertrial interval that are not differentially reinforced by the experimenter are frequently referred to as 'spontaneous' or intertrial responses. Sheffield (1965) referred to them as an "ever-present nuisance," implying that they make it less easy to determine the precise extent of conditioning. Most investigators either ignore, or expend considerable effort to minimize, the occurrence of those responses, but little is known about the factors responsible for eliciting them.

In an attempt to make theoretical sense of intertrial responses, Mowrer and Lamoreaux (1951) suggested that they simply reflect the ongoing process of discrimination learning involved in a conditioning procedure. Such a notion implies that the intertrial responses are generalized avoidance responses elicited by the fear associated to the background cues of the apparatus. The extent of this generalization may be considered a function of the difficulty the subject has distinguishing situations with the conditioned stimulus (the CS, in the CS-US interval) from situations without the CS (intertrial interval). It would be expected, then, that intertrial responding should increase early in training, when that distinction is low, and decrease as the conditioned avoidance response, and the discrimination of the CS, becomes well established—a *trials* effect. Such tends to be the general finding in the literature (Black and Carlson, 1959; Brush, 1957; Church, Brush, and Solomon, 1956; Kamin, 1954; Thompson, Sachson, and Higgins, 1969).

Another prediction that would follow from Mowrer's generalized-response hypothesis is that intertrial responses should proportionally decrease as the relative stimulus difference between the CS-US and intertrial intervals increases. One method to increase this difference is to increase the *complexity* of the CS by adding components. Although the role of CS complexity in the shuttlebox avoidance responding has been investigated (Levis, 1970; Levis and Stampfl, in press), an in-depth analysis of its effects on intertrial responding had yet to be made. So the present study compared the effects of two different complexities of nonserial CS presentation and two different complexities of serial CS presentation on the development of intertrial responding in a shuttlebox situation with CS-contingent (paired) shock. The effects of CS-separate shock and of no shock were also evaluated.

METHOD

—Subjects—The subjects were 200 experimentally naive male Sprague-Dawley rats obtained from commercial breeders. They ranged in age from 90 to 110 days.

—Apparatus—The apparatus, constructed of ⅜-in. clear Plexiglas, was 36 in. long, 5 in. wide, and 10 in. high. The box was divided into two 18-in. compartments by a ⅛-in. Masonite barrier extending 4 in. above the surface of the floor. The floor of each compartment consisted of a grid made of ⅛-in. brass welding rods spaced ½ in. apart, center to center. A response was defined as a jump over the barrier from one compartment to the floor of the other. Each compartment of the apparatus could be independently programmed to activate: three pairs of 7½-w, 115-v screw-base lamps that flashed twice per second; a tone; and a shock source. The tone, a 1,275-Hz sine wave, was produced by an audio signal generator. A Grason-Stadler shock source with scrambler was set to produce 1 ma of current. Avoidance, escape latencies, and the frequency of intertrial responding were recorded by a Grason-Stadler print-out counter (E 12505A).

—Experimental groups—A factorial design (four types of CS presentation, two levels of CS order) was used. Forty rats were assigned unsystematically to each presentation condition.

Two of the *types* of presentation (one-component and two-component compound) involved nonserial manipulations, and two (two-component serial and two-component serial compound) involved serial manipulations. For the nonserial manipulations, only one of the manipulated CSs was present for the one-component condition (S_1) during the designated CS-US intervals, whereas two CSs were presented simultaneously throughout the CS-US interval for the two-component compound condition (S_1S_2). For both of the serial manipulations two CSs were presented, with the onset of S_2 occurring at the midpoint of the CS-US interval and dividing the interval into two equal time segments. For the serial condition (S_1/S_2), S_1 was terminated when S_2 was initiated, while in the serial compound condition (S_1/S_1S_2), S_1 continued to remain on for the entire CS-US interval.

In group S_1, the *order* of presentation was simple: a tone served as CS for half the rats; flashing lights as CS for the others. In group S_1S_2, all rats were simultaneously presented both tone and lights as CS. In group S_1/S_2, half the rats got the lights followed by the tone as CS, and half got the tone followed by the lights. In group S_1/S_1S_2, half the rats got the lights followed by a combination of lights and tone, and half got the tone first and then the tone plus lights.

Each rat was given a 30-min adaptation period to the apparatus before the start of the first trial. A delayed-conditioning procedure with a CS-US interval of 16 sec was used. An escape response terminated both CS and US; an avoidance response terminated CS and prevented US onset. Each rat was given 150 consecutive acquisition trials in a single session. The length of the intertrial interval averaged 60 sec (50–60–70 sec, randomized). The order of running was determined randomly.

—Control groups—Two groups of 20 rats each were used, one with CS and shock presented separately and one with CS but no shock presented. The serial compound CS (S_1/S_1S_2) was used for both control groups. Group NS got no shock during their trials. Group SS received shock an average 30 sec (20–30–40 sec, randomized) after termination of the CS. If the rat failed to respond, the CS was terminated 20 sec after its onset, a 20-sec interval being used because the CS-contingent shock for the experimental groups resulted in a CS duration longer than 16 sec on escape trials (S_2 still occurred at 8 sec unless a response latency was shorter than 8 sec, however). All other aspects of the procedure for control rats were the same as those for experimental rats. (For a more detailed presentation of the apparatus and procedure of this experiment, see Levis and Stampfl, Experiment I, in press.)

RESULTS AND DISCUSSION

The mean numbers of conditioned avoidance responses for each group and the corresponding numbers of intertrial responses are shown in Table 1. Comparisons between groups S_1 and S_1S_2 and between groups S_1/S_2 and S_1/S_1S_2 showed no reliable differences on total numbers of avoidances. As expected, however, an analysis over the total numbers of intertrial responses revealed an effect of CS complexity: group S_1 differed significantly from group S_1S_2 ($p < .01$) and group S_1/S_2 from group S_1/S_1S_2 ($p < .025$), with the latter, more complex, CS in each comparison producing fewer intertrial responses. Groups S_1S_2 and S_1/S_1S_2, which produced the smallest numbers of intertrial responses, were the only experimental groups that had two stimulus components present simultaneously for at least half the CS-US interval.

Separate analyses of the intertrial-response data within each experimental group in terms of the order of the CS revealed only one reliable

Table 1. Mean avoidance and intertrial responding

	Group S_1		Group S_1S_2	Group S_1/S_2		Group S_1/S_1S_2		Group NS		Group SS	
	T	L	TL	L/T	T/L	L/LT	T/TL	L/LT	T/TL	L/LT	T/TL
Avoidance responses	53.0	53.5	53.5	57.4	70.6	77.0	76.8	4.6	.5	2.7	1.6
Intertrial responses	41.5	25.5	9.9	16.0	27.4	3.5	11.2	1.3	.0	5.6	1.1
Learners	45.3		18.3	38.6		10.1					
Nonlearners	2.8		1.0	1.1		7.3					

Note: T (tone) and L (flashing lights) refer to the order of CS presentation in each group and subgroup. The data for the learners and nonlearners combine the orders within each experimental group.

difference. That was for group S_1/S_1S_2, where the rats that got the tone first and then the tone plus flashing lights produced significantly more intertrial responses than the rats that got the lights first and then the combination of lights and tone ($F = 5.05$, $df = 1/38$, $p < .05$).

An analysis of the total intertrial responses across blocks of 10 trials revealed no significant trials effect nor any significant interaction of trial and group. This failure to obtain a trials effect calls into question Mowrer's discriminative notion, outlined in the introduction. An alternative hypothesis might be that intertrial responses are only random jumps, completely independent of CS-US pairing. Groups NS and SS are the appropriate controls for an evaluation of that alternative, and as Table 1 shows, both groups made only occasional responses. Clearly, shock in and of itself did not facilitate intertrial responses, nor did it differentially suppress their rate, since the responding in groups NS and SS was not reliably different. But when a comparison between those control groups and group S_1/S_1S_2 was made (all three groups having the same CS and the same order of presentation), there was a reliable difference in intertrial responding, which clearly suggests that the contingency of CS-US pairing *was* a critical factor in eliciting intertrial responding.

Finally, the number of the 'spontaneous' responses made during the 30-min adaptation period for each rat in the experimental groups was correlated with the trial number of the rat's first conditioned avoidance response ($r = -.09$) and with its total number of conditioned avoidance responses ($r = +.02$). In brief, no reliable effect appeared.

Learners versus nonlearners

It seemed possible that 'poor conditioners' in the experimental groups might have been responsible for the nonsignificant trials effect. To check this possibility, the rats were divided into 'learners' (animals that reached a criterion of 10 consecutive avoidances within 150 trials) and 'nonlearners.' The number meeting the learning criterion across experimental conditions ranged from 20 to 28; the number failing to meet it ranged from 12 to 20. For purposes of analysis (see Spence and Platt, 1966), the 20 learners that produced the highest mean number of intertrial responses and the 12 nonlearners that yielded the lowest number were included. Table 1 compares the mean number of intertrial responses for each experimental group separately for learners and nonlearners. This learner-nonlearner breakdown was extended to the selected animals across blocks of 10 trials.

The comparison between learners and nonlearners yielded a significant difference ($F = 19.71$, $df = 1/158$, $p < .01$), the mean number of intertrial responses across all 150 trials being 28.1 and 3.1 for learners and nonlearners respectively. And as expected, both the trials effect ($F = 2.69$, $df = 14/1764$, $p < .01$) and the interaction of that and learners versus nonlearners ($F = 2.36$, $df = 14/1764$, $p < .01$) were significant.

A comparison of the learners in each of the four experimental groups across trial blocks also produced a significant group effect ($F = 4.15$, $df = 3/76$, $p < .01$). Group S_1 differed from group S_1S_2 ($p < .05$), and group S_1/S_2 from group S_1/S_1S_2 ($p < .01$), with the latter group in each comparison producing fewer intertrial responses. The desired trials effect for learners was also obtained ($F = 3.14$, $df = 14/1064$, $p < .01$) and is graphically presented in Figure 1 along with the nonsignificant trials effects for the nonlearners. It should be noted, however, that the function graphed for learners is most likely flattened because of the varying acquisition rates for individual rats. The interaction between trials and groups was not significant. A similar analysis over the nonlearners did not yield any reliable differences. (The higher mean intertrial responding for the nonlearners in group S_1/S_1S_2 seemed attributable to one animal.)

Because the distribution of intertrial responding over the course of avoidance conditioning was of interest, a Vincent analysis was performed on both the avoidance- and intertrial-response data of the top

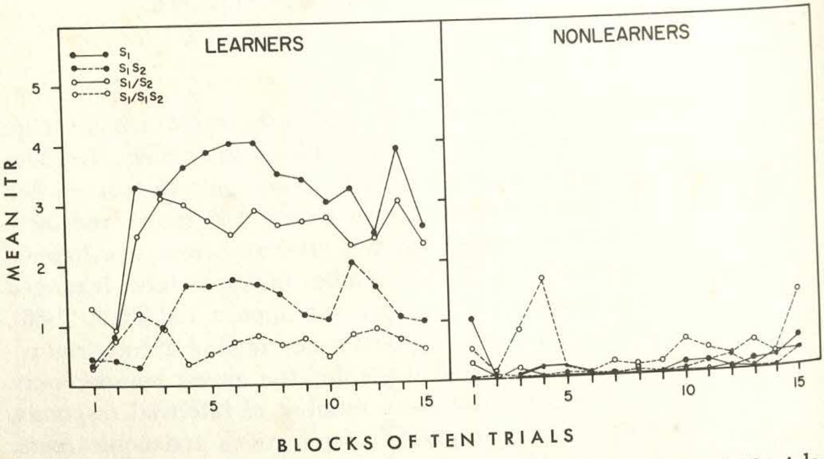


Fig. 1. Mean numbers of intertrial responses (ITR) across blocks of 10 trials; data plotted separately for learners and nonlearners of each experimental group

20 learners in each experimental group. Accordingly, trials through criterion for each rat were divided into seven equal trial blocks[1] and the per-trial rates of avoidance and intertrial responding were computed for each block for each animal. The mean for each block was then converted into a percentage of total responding and plotted for both sets of response data in Figure 2. It is apparent from this figure that the percentages of avoidance and intertrial responses increased at approximately the same rate during acquisition—or to put it in other words, that intertrial responding correlated positively with conditioned avoidance responding across these trials. An inspection of the postacquisition data further indicated that the number of avoidance responses for each experimental group became relatively stable across these trials while the intertrial responses tended to show a slight decline.

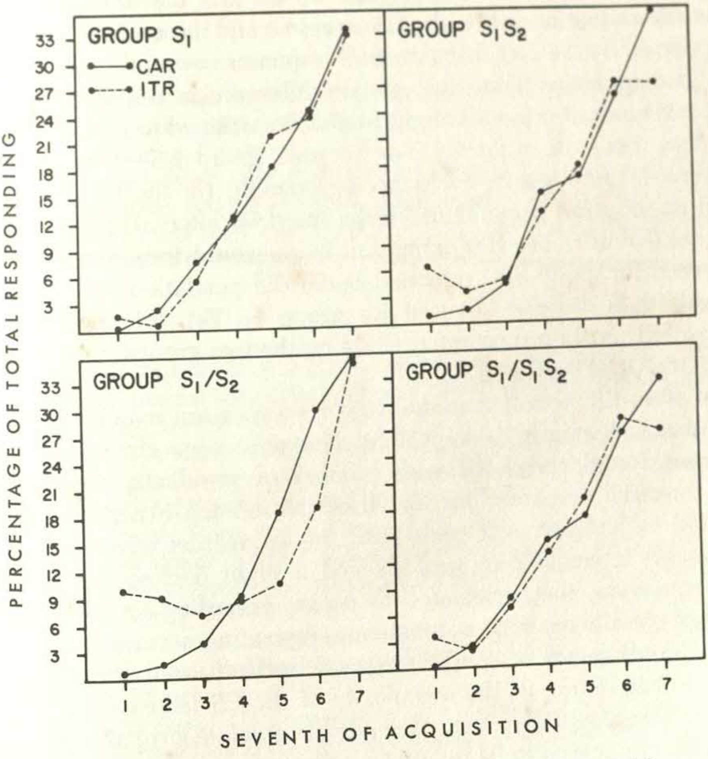


Fig. 2. Mean conditioned avoidance responses (CAR) and intertrial responses (ITR) as percentages of total responding for each seventh of acquisition; data plotted for learners in each experimental group

Another aspect of the data that is quite informative is the finding that the learners in groups S_1 versus S_1S_2 and groups S_1/S_2 versus S_1/S_1S_2 did not show reliable differences in total number of avoidance responses, even though they did show reliable differences in intertrial responses. Subsequent comparisons of these groups on trial number of first avoidance and on longest consecutive series of avoidances also produced no reliable differences. So although, as Figure 2 demonstrates, a high frequency of intertrial responses is *associated* with the acquisition of avoidance responses, the frequency of that intertrial responding is not necessarily directly related to the frequency of avoidance responding. In fact, the absolute level of intertrial responses apparently had no direct bearing on the absolute level of avoidance responses. That is to say, intertrial responses produced neither a facilitory or inhibitory effect.

The assumption that intertrial responses are directly related to a lack of differentiation between the CS-US and the intertrial intervals (between situations when the CS is present and those when it is not) was supported by the fact that intertrial responses covaried with the degree of CS complexity. Thus, the relative difference in responding between the two kinds of intervals might further be assumed to provide an 'index of discrimination' of the CS. For example, group S_1S_2 showed a greater difference between its avoidance responding (in the CS-US interval) and its intertrial responding (in the intertrial interval) than did group S_1; see Table 1. The discrimination index would thus suggest that the CS used for group S_1S_2 provided better differentiation between the intervals than did the CS used for group S_1. Yet, this 'differentiation' apparently had no substantial effect on the two groups' avoidance performance; also see Table 1.

In sum, the notion that intertrial responses are mainly generalized avoidance responses is appealing. Not only were their intertrial responses found correlated with number of avoidance responses, but the intertrial responses for the 'learners' in each experimental group tended to increase over trials until the acquisition criterion was met. Although intertrial responses showed a slight decline past the point of acquisition, they continued to occur. According to data from the control conditions, little spontaneous responding occurred unless there was a contingency of conditioned and unconditioned stimuli (CS and US). Furthermore, as the complexity of the CS increased, the rate of intertrial responding decreased, while the absolute level of avoidance responding appeared to be relatively independent of the absolute level of intertrial responding.

Additional data were obtained with a different strain of rats (Iowa hoods). A one-component CS (S_1) was compared with a three-component compound CS ($S_1S_2S_3$), and a two-component serial compound (S_1/S_1S_2) was compared with a three-component serial compound ($S_1/S_1S_2/S_1S_2S_3$), using apparatus and procedure similar to those just reported. These replication data were quite consistent with the main findings. Intertrial responding correlated positively with avoidance responding. It increased with trials and subsequently decreased and leveled off. Learners differed from nonlearners, and the discrimination index was significantly larger for the CS, in each comparison, that encompassed the greater number of components. So the main results of the present study appear to be independent of the strain used. We might finally note that, although more direct comparisons are needed, hooded rats tended to produce less intertrial responding.

Notes

The research reported was supported by Biomedical Science Support Grant FR 07035 and by Grant MH 16584 from the National Institutes of Health. Received for publication October 19, 1970.

1. The trials remaining after dividing by 7 were randomly distributed among the seven blocks with two restrictions: that no block for any subject have more than one 'extra' trial, and that the extra trials be evenly distributed over blocks.

References

Black, A. H., and Carlson, N. J. 1959. Traumatic avoidance learning: A note on intertrial-interval responding. *Journal of Comparative and Physiological Psychology* 52:759–760.

Brush, E. 1957. Traumatic avoidance learning: The effects of conditioned stimulus length in a free-responding situation. *Journal of Comparative and Physiological Psychology* 50:541–546.

Church, R. M., Brush, F. R., and Solomon, R. L. 1956. Traumatic avoidance learning: The effects of CS-US interval with a delayed-conditioning procedure in a free responding situation. *Journal of Comparative and Physiological Psychology* 49:301–308.

Kamin, L. J. 1954. Traumatic avoidance learning: The effects of CS-US interval with a trace-conditioning procedure. *Journal of Comparative and Physiological Psychology* 47:65–72.

Levis, D. J. 1970. Serial CS presentation and shuttlebox avoidance conditioning: A further look at the tendency to delay responding. *Psychonomic Science* 20:145–147.

Levis, D. J., and Stampfl, T. G. 1971. Effects of serial CS presentation on shuttlebox avoidance responding. *Learning and Motivation,* in press.

Mower, O. H., and Lamoreaux, R. R. 1951. Conditioning and conditionality (discrimination). *Psychological Review* 58:196–212.

Sheffield, F. D. 1965. Relation between classical conditioning and instrumental learning. In *Classical conditioning: A symposium,* ed. W. F. Prokasy. New York: Appleton-Century-Crofts.

Spence, K. W., and Platt, J. R. 1966. UCS intensity and performance in eyelid conditioning. *Psychological Bulletin* 65:1–10.

Thompson, C. P., Sachson, S. M., and Higgins, R. L. 1969. Distribution of intertrial responses in shuttlebox avoidance conditioning. *Journal of Comparative and Physiological Psychology* 69:563–572.

American Journal of Psychology
1971, Vol. 84, No. 4

A SIZE-CONTRAST ILLUSION WITHOUT PHYSICAL SIZE DIFFERENCES

Stanley Coren
New School for Social Research

A stereogram of equally sized circles produced a size-contrast illusion in 11 of 12 observers, providing evidence that the illusion is of central origin and occurs sometime after the operation of the constancy mechanism.

One of the size-contrast illusions best known is the Ebbinghaus figure (frequently referred to as Titchner's circles). In this figure a circle surrounded by small circles is seen as larger than a circle surrounded by large circles. Although frequently reproduced in elementary texts, very little work has actually been done on this configuration, and an adequate explanation of its effect is still lacking.

Ohwaki (1960), Springbett (1961), Day (1961), and Schiller and Wiener (1962) have shown, however, that the illusion is probably not of retinal origin, since the test and inducing elements can be placed into different eyes and the illusion still seen in the fused image. Further evidence that central processes may be involved is found in the demonstration shown in Figure 1. When stereoscopically fused the top configuration appears as a central disc with four surrounding circles imaged some distance behind the plane of the central test circle; the lower configuration has the central disc at the same stereoscopic distance as the test disc in the first group, but the surrounding circles are seen as considerably closer. By the operation of size constancy the surrounding circles that are apparently closer appear smaller than the surrounding circles that are apparently farther away, although all of the circles in the figure are physically the same size.

The relevant question has to do with the apparent size of the central circles in Figure 1, those circles which are stereoscopically at the same distance. Twelve observers were used, all with normal or corrected visual acuity of 20/25 or better and with normal stereopsis as measured on a Keystone telebinocular. They viewed the figures through a lens stereoscope and were asked to judge the relative size of the central discs in a forced-choice paradigm. After the first judgment, the left and right views were interchanged so as to reverse the disparities and the

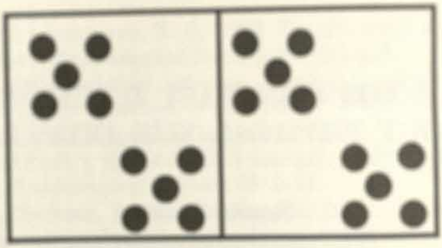


Fig. 1. A stereogram that produces a size-contrast illusion without physical size differences

subjects were required again to judge the relative size of the central circles. In 11 out of the 12 observers, the central disc surrounded by the apparently larger circles appeared smaller than the central disc surrounded by the apparently smaller circles. This is significant, with $p < .01$ by a sign test.

This finding would seem to indicate that this size-contrast illusion is not only of central origin but occurs sometime after the operation of the constancy mechanism. Although there is no immediate explanation for the phenomenon, it is interesting that Gregory's theory of illusions (1966) requires that illusionary distortions occur at the same level or after the operation of constancy scaling. This demonstration indicates that size-contrast illusions may occur *beyond* that level.

Notes

Received for publication December 3, 1970.

References

Day, R. H. 1961. On the stereoscopic observation of geometric illusions. *Perceptual and Motor Skills* 13:247–258.
Gregory, R. L. 1966. *Eye and brain*. New York: McGraw-Hill.
Ohwaki, S. 1960. On the destruction of geometrical illusions in stereoscopic observation. *Tohoku Psychologica Folia* 29:24–36.
Schiller, P., and Wiener, M. 1962. Binocular and stereoscopic viewing of geometric illusions. *Perceptual and Motor Skills* 15:739–747.
Springbett, B. M. 1961. Some stereoscopic phenomena and their implications. *British Journal of Psychology* 52:105–109.

American Journal of Psychology
1971, Vol. 84, No. 4

AN INSTRUMENTAL-LEARNING TASK FOR HUMAN SUBJECTS

In-mao Liu and Chor-ping Hui
National Taiwan University

The task proposed involves a space in which a chain of instrumental acts takes place, and a program depicting the nature of those acts. The space and the program are independently manipulatable, and both affect instrumental learning. Experiments I and II study the properties of the space; Experiments III and IV, the nature of the program.

Most instrumental-learning situations are devised for animals. Some of them are directly applicable to humans, such as stylus mazes derived from ordinary mazes. In this case, however, the subjects have to be blindfolded. The mirror star-tracing task can be also an instrumental one, but it seems to require an unusual skill.

There are also serial-search tasks, in which the finding of one clue gives directions for finding the second, the second for the third, and so on. The Tsai-Partington Numbers Test (Ammons, 1955) is an example. In this test a series of numbers, from 2 through some large number like 20, for instance, are randomly scattered on a sheet of paper. In the middle of the paper, the digit 1 appears, distinctly set off from the other numbers. Given a signal to start, the subject places his pencil on the 1, then searches for 2 and draws a line to it, then searches for 3, draws a line from 2 to 3, and so on. This task permits the experimenter to study separately each act of reaching the succeeding number from a given number.

From the standpoint of experimental analysis, however, it is desirable to have a task in which each required act can be manipulated independent of the stimulus characteristics of the task situation. And in the cases of search tasks such as one developed by Bakan (1955), as well as the Numbers Test described above, each act depends entirely on the stimulus characteristics of the task situation.

In the present paper, a new learning task for human subjects is described. This task consists of a space and a program. The space is a matrix of points. The program is made up of a certain number of directions, each direction telling the subject how to move from a point in the space to the next. This task enables the experimenter to manipulate the characteristics of the space (within which a chain of acts takes place) separately from the program (which specifies the nature of that chain).

THE TASK ITSELF

A space consisting of $n \times n$ matrix points is printed on a sheet of paper. Two adjacent points are always separated by a distance of 1 cm. (The point of the *i*th row and *j*th column is henceforth referred to as point *i,j*.) The subject's task is to place his pencil on a marked starting point in the space and then consecutively to draw a circle around each of the correct points in the space by following each of a consecutive sequence of directions—the program.

There were several types of directions used in the present experiments. For the direction ⟷, the subject was required to move from a given point to the point *symmetrical* with it on the vertical axis bisecting the space. For instance, starting from point 2,3, with this direction the subject was to move to point 2,6 in the space of 8×8 matrix points. The directions $\updownarrow$ and $\nearrow$ (or $\nwarrow\!\!\searrow$) were similarly defined; they carried the subject from point 2,3 to points 7,3 and 3,2 (or 6,7) respectively. For the *nonsymmetry* directions $\xrightarrow{2}$ (or $\xleftarrow{2}$, $2\uparrow$, $2\downarrow$), $3\nearrow$ (or $3\swarrow$, $\nwarrow 3$, $\searrow 3$), and $\underset{\downarrow}{\overset{1}{\neg}}2$ (or $\overset{\uparrow}{\underset{1}{\neg}}\,2$, $2\overset{\uparrow}{\underset{1}{\neg}}$, $2\underset{\downarrow}{\overset{1}{\neg}}$), the subject was respectively to move to the point two steps away, to the point three steps in the indicated diagonal direction, and to the point one step away horizontally plus two steps away in the vertical direction.

The experimenter first explained the meaning of the various types of directions in the programs to the subject. On each trial, the subject had a sheet of paper on which a space of $n \times n$ matrix points was printed. He was instructed in advance to place his pencil on the marked starting point, then to move to successive points by following the program of directions. When he made an error by drawing a circle around a wrong point, he was informed immediately and was asked to move his pencil back to the last correct point to continue anew. As soon as his pencil reached the starting point, a clock started; it was stopped when he drew a circle around final point of the program. Figure 1 shows the correct moves for a subject following the program of Experiment I.

EXPERIMENT I

By analogy from maze situations, it is easily predicted that subjects should take more time to complete a longer program than a shorter one.

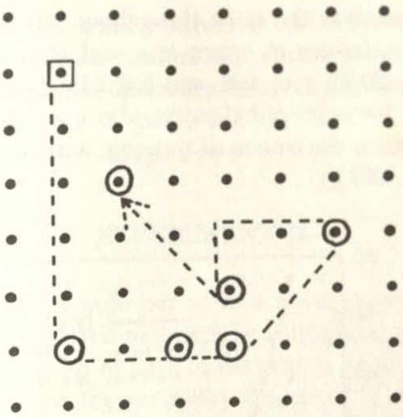


Fig. 1. Execution of the program of Experiment I (↕, $\xrightarrow{2}$, ⟷, 2 ↗, 1 $\overset{2}{\swarrow}$, ⤡) in the space of 8 × 8. A point with a square around it is the starting point. The dotted line indicates the path of carrying out the program.

However, the space in which they complete a program should be a more interesting variable to study. With the same program, it might take them longer time to complete the task in the space of 16 × 16 than in the space of 8 × 8. Experiment I investigated this '*size*' property of the space.

Method

—Subjects and task—The subjects were 30 juniors in the department of psychology; they participated in the experiment to fulfill part of a course requirement. They were divided into three groups according to order of their appearance, each group being assigned to one of the three spaces: 4 × 4, 8 × 8, or 16 × 16. The program used for all three groups consisted of six directions: ↕, $\xrightarrow{2}$, ↔, 2 ↗, 1$\overset{2}{\swarrow}$, and ⤡. The starting point was always point 2,2. Each subject was run 20 massed trials.

Results and discussion

Figure 2 shows the mean time to complete the program for each trial block for each group. Irrespective of difference in the number of points in the spaces on which the subjects worked, it took about 20 trials to reach asymptote; there was practically no sign of improvement during the last three trials. At the beginning of training, the three groups differed considerably in such a way that the subjects working with the 'larger' spaces took much more time to complete the program. As training proceeded, the difference in time became smaller, although the direction of difference was the same throughout. An analysis of variance showed that the variables of space size and training had significant effects [$F(2, 27) = 29.21, p < .001$; and $F(9, 243) = 235.61, p < .001$, respectively]. Their interaction, indicative of convergence of the absolute magnitudes of time in the course of training, was also significant [$F(18, 243) = 13.24, p < .001$].

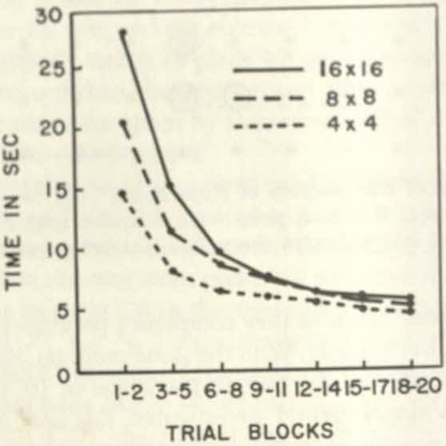


Fig. 2. Acquisition curves showing the effect of space size

A preliminary study had shown that the length of path (as opposed to length of program) was not an important determinant of the time required to complete a program. In the present case, the lengths of path required to complete the program in the spaces of 4 × 4, 8 × 8, and 16 × 16 were 14.6, 15.8, and 49.4 cm respectively. But although the difference in path lengths for the spaces of 4 × 4 and 8 × 8 was negligibly small, there was a reliable difference in time required to complete the program between these two groups.

The obtained differences in time to complete the program must be attributed to the difference in carrying out the directions ↕, ↔ , and ↘ in the different spaces. In carrying out the instrumental act required for each of these symmetry directions, the subject had to select the correct point from a greater number of points when he was working in the larger spaces. In other words, the larger amount of information was associated with the larger space. Taking performance levels on the last trial block as asymptotes, it is apparent from Figure 2 that the time required to complete the same program in the different spaces was approximately linear with the amount of information associated with the spaces. This is because the difference between the amounts of information in the 4 × 4 and 8 × 8 spaces was the same as the difference between the 8 × 8 and 16 × 16 spaces.

EXPERIMENT II

Even if the length of path per se in a space is not important in determining the time required to complete a program, some other characteristics, such as the *locus* of path in that space, might be. For example, if the starting point for the symmetry directions is on the margin of a space, its destination point should be also on the opposite margin of the space, as that would undoubtedly (it seemed) facilitate the subjects' selection of the correct destination point. The hypothesis, in brief, was that the locus of path and the overall information in a space might both be factors in determining the time required to complete a program.

Consider an 8 × 8 space. This space (L) can be divided into two subspaces: the first (L_1) consisting of all those points (i,j) such that i or $j = 1, 2, 7$, or 8; the second (L_2) consisting of the rest of the points. In this way, the second subspace consists of the inner 4 × 4 space by deleting the first and last two rows and two columns of the original 8 × 8 space.

Obviously, if the starting and all successive points of a program are within L_1 subspace, the time required to complete this program should be shorter than a program whose starting and successive points are all within L_2 subspace—if it is indeed more difficult to select middle points than marginal points. Secondly, by the same logic, if the starting and successive points of a program are all within L_2 subspace, the time required to complete this program should be considerably longer than the time required to complete the same program when L_2 subspace is presented alone, even though the lengths of path in the two conditions are exactly the same. In other words, Experiment II was also designed to see whether the locus of path alone is an important variable in determining the time required to complete a program.

Method

—Subjects and tasks—The subjects were 12 graduate students in the department of psychology. They voluntarily participated in the experiment.

The conditions of this experiment are presented in Table 1. In condition L_1 each subject was required to carry out a program in the L space of 8 × 8. However, the starting and successive destination points were all within L_1 subspace. Each subject in condition L_2 was presented the same L space, but the starting and destination points were all within L_2 subspace. In condition L_2-alone each subject was presented the L_2 space of 4 × 4 alone. For each condition, two different programs were prepared.

—Procedure—Each subject worked on a program in each condition. The selection of either of the two programs for each condition and their ordering in the three conditions were counterbalanced between subjects. For each program, each subject worked for 20 trials.

Since this is a within-subjects design, the same program requiring the same locus of path was not used under conditions L_2 and L_2-alone. However, the programs were so chosen that their lengths of path were approximately equivalent (see the averages for path length in Table 1).

Table 1. Design for Experiment II

Condition	Starting point	Program directions 1	2	3	4	5	6	Path length (cm)	Average path length (cm)
L_1 in	2,2	↕	$\overset{1}{↰}{}_{2}$	↔	$^{2}↙$	$\overset{2}{\leftarrow}$	⤡	22.6	20.9
8 × 8	1,7	$^{2}\downarrow$	↔	⤢	↕	$^{1}\underset{2}{↳}$	$^{2}↗$	19.2	
L_2 in	3,3	↕	$\underset{2}{\lrcorner}\uparrow^{1}$	↔	$^{2}↗$	$^{2}\downarrow$	⤡	14.6	15.6
8 × 8	3,4	⤡	$\overset{2}{\leftarrow}$	$^{2}↗$	↕	↔	${}_{1}\overset{2}{↱}$	16.6	
L_2-alone	1,2	$↘^{2}$	$^{1}\underset{2}{\uparrow\llcorner}$	⤡	↕	$^{2}\downarrow$	↔	11.2	14.7
of 4 × 4	2,3	${}_{1}\overset{2}{\downarrow\ulcorner}$	↔	$↖^{2}$	$\overset{2}{\rightarrow}$	⤢	↕	18.1	

Results and discussion

In the course of training within each condition, performance under condition L_2-alone was found to be consistently superior to that under condition L_2, and performance under condition L_1 was also superior to that under condition L_2. An analysis of variance showed that the condition, training, and their interaction were all significant sources of variance [$F(2, 22) = 112.34$, $p < .001$; $F(9, 99) = 229.41$, $p < .001$; and $F(18, 198) = 9.28$, $p < .001$, respectively]. The significant interaction illustrates again the convergence phenomenon found in Experiment I.

Considering performance on the last three trials as at asymptotic levels, this performance level under condition L_2-alone was superior to those under conditions L_2 and L_1 [$t(11) = 7.17$, $p < .001$; and $t(11) = 3.15$, $p < .01$, respectively]. In addition, asymptotic performance under condition L_1 was superior to that under condition L_2 [$t(11) = 4.31$, $p < .01$]. Specifically, the mean asymptotic performance (in seconds) and the standard deviation for the L_1 subjects were 4.27 and 1.11 respectively, for the L_2 subjects they were 5.26 and 1.29, and for the L_2-alone subjects they were 3.83 and 1.11.

Although the length of path of a program in condition L_1 was systematically longer than in condition L_2, the superior performance under condition L_1 compared to that under condition L_2 must be due to the locus of path, since performance under condition L_2-alone was superior to that under condition L_2 even though the lengths of path were approximately equal in these last two conditions.

Since each subject was given three programs in Experiment II, it was also possible to obtain some index of reliability of the present task. Using asymptotic performance as the response measure, the correlations between conditions L_1 and L_2, between conditions L_2 and L_2-alone, and between conditions L_2-alone and L_1 were respectively .79, .84, and .90. Therefore, the present task is quite reliable. It is to be noted that this high reliability is based on asymptotic performance after intensive training on each condition. Practice effects usually showed up for the first few trials when the condition was changed from one to another.

EXPERIMENT III

Experiment III attempted to explore the nature of the group of instrumental acts defined by the program. The question here was whether the acts depend entirely on the program *directions* per se (their sequence) or on the starting and successive points (the *path*) those directions describe. If the acts defined by a program depended completely on the program's starting point, for example, the same program with different starting points could define entirely different groups of instrumental acts.

For this purpose, three experimental conditions sufficed. In the path-relevant condition, each subject carried out the same program from the same starting point for a given number of trials. In the path-irrelevant condition, each subject carried out the same program with different starting points on successive trials. In the control condition, each subject carried out different programs from different starting points on successive trials. If the performance under the path-irrelevant condition did not differ from that under the path-relevant condition, we would know that a program completely defines a group of instrumental acts without reference to its starting point (or here, its path). But if the performance under the path-irrelevant condition lay between those under the other two conditions, we would know that in defining a group of instrumental acts, a program partially depends on its starting point (or here again, its path).

Method

—Subjects—The subjects were 48 sophomores and juniors in the department of psychology. They participated in the experiment voluntarily. They were divided into three groups according to order of their appearance, with the sexes equally represented in the groups.

—Tasks and procedure—For one group (path-relevant condition), each subject was given the same program with a fixed starting point on 25 successive trials. Two programs were prepared, so that half of this group worked with one program and the other half with the other. For another group (path-irrelevant condition), each subject was given the same program with varied starting points on 25 successive trials. Two programs were used, as in the path-relevant condition. For still another group (control condition), each subject was given 25 different programs whose starting points differed from one another, on 25 successive trials. All the programs in Experiment III consisted of six directions, and the space of 8 × 8 was used.

Results and discussion

Again, the average of the last three trials was obtained for each subject and used as asymptotic performance level. The means of those averages and the standard deviations were 4.68 and .94 respectively for the path-relevant subjects, 10.19 and 1.99 for the path-irrelevant subjects, and 13.77 and 2.33 for the control subjects. Although performance under the path-irrelevant condition fell between those under the two other conditions, it was closer to the control condition.

An analysis of variance showed that asymptotic performance was significantly affected by experimental condition [$F(2, 46) = 43.16, p < .001$]. Further analyses showed that performance under the path-relevant condition was superior to that under the path-irrelevant and control conditions, and that performance under the path-irrelevant condition was superior to that under the control condition [$t(46) = 5.62, p < .001$; $t(46) = 9.28, p < .001$; and $t(46) = 3.65, p < .001$, respectively].

The improvement in performance with varying programs and varying starting points (control condition) shows the so-called learning to learn. Since this is a well-known phenomenon, no particular statistical analysis was made. However, the fact that performance under the path-irrelevant condition fell between those under the other two conditions demonstrates yet another phenomenon—that in defining a group of instrumental acts a program depends partially on its starting point or, equivalently here, on the path it describes. With repeated trials in the path-relevant condition, the subjects apparently learned the program as a group of directions as well as by its path. When the path was experimentally separated from the directions used to define the program (path-irrelevant condition), performance was inferior to that in the path-relevant condition, where the subjects could also rely on the specific path to help them carry out the program. Nevertheless, since performance under the path-irrelevant condition was superior to that under the control condition, the subjects also learned the program as a group of directions independent of its path.

EXPERIMENT IV

Experiment IV was also concerned with the nature of the group of instrumental acts defined by the program. The main problem here, however, had to do with the *similarity* of groups, or chains, of instrumental acts. Consider several programs each of which consists of six directions. A given program certainly seems more similar to a program whose initial five directions are the same as its own than to a program whose initial four directions only are the same. An analogous statement should apply (we hypothesized) to a collection of programs whose initial segments of directions vary from one another while the last parts of directions are the same.

In view of the results of Experiment III (namely, that the group of instrumental acts defined by a program depends on the sequence of directions as well as on its path), Experiment IV attempted to study two types of similarity among a collection of groups of instrumental acts. For one type of similarity, the standard group (chain) of instrumental acts was defined by a fixed sequence of directions with a fixed starting point. The first or last parts of the sequence of directions were then systematically varied for comparison with the standard, the chains thus constructed being thereby equivalent to the test stimuli in experiments on stimulus generalization. For another type of similarity, the standard chain of instrumental acts was defined by a fixed sequence of directions with varying starting points. In other words, the path with these latter chains was not fixed. As in the case of the first type of similarity, however, various chains of instrumental acts were generated on the test trials by systematically varying the first and last parts of the sequence of directions.

Method

—Subjects—The subjects were 28 students in the department of psychology at Chung Yuan Christian College of Science and Engineering. They served in the experiment voluntarily.

—Procedure—The subjects were divided into two groups according to order of their appearance. For one group (group FPS), the training trials had each subject working on a fixed program with a fixed starting point. On the test trials, which were interspersed among the last part of the training trials, the following five types of programs were administered once each: (*a*) a program whose last direction differed from the standard program; (*b*) a program whose last two directions differed from the standard; (*c*) a program whose first direction differed from the standard, implying that the starting point differed from the standard and that, after carrying the first direction, the remaining directions as well as the remaining path were the same as the standard; (*d*) a program whose first two directions differed from the standard; and (*e*) an entirely different program with a different starting point.

For another group (group FP), each subject worked on a fixed program with varying starting points on the successive training trials. There were also five types of test trials, corresponding to those for group FPS. The only difference was that there was no standard fixed path.

For both groups the first 25 trials were the training ones with the standard program. Then, the five test trials interspersed among further training trials were administered, their order systematically varied between subjects. The numbers of training trials inserted between two consecutive test trials were 4, 2, 3, and 3. Thus for each subject 37 training and 5 test trials were run. For each group, there were two standard programs and corresponding sets of test programs. Half of the subjects in each group received one standard program and its accompanying test programs; the other half, another standard program with its accompanying tests. All the programs consisted of six directions, and the space of 8 × 8 was used throughout.

Results and discussion

On the question of *path* similarity, it is the data for group FPS that are relevant. The upper part of Table 2 shows the forward- and backward-generalization gradients obtained when the path described by the standard program was systematically varied on the test trials (by methods *a* through *e* above); the time it took to complete the standard program was obtained by averaging the last three training trials before administering the first test program for each subject. The entries in the row for the forward-generalization gradient represent those cases in which the initial path was the same as the standard and the remaining path was different (F_1 and F_2); the entries for the backward-generalization gradient, those cases in which the final path was the same and the initial path was different (B_1 and B_2).

Both forward- and backward-generalization gradients showed that the time required to complete a test program increased as the test increasingly differed from the standard. Separate analyses of variance showed that paths (here expressed as programs) varying in their similarity were significant sources of variance [$F(3, 39) = 36.51, p < .001$; $F(3, 39) = 48.80, p < .001$] for both forward and backward cases respectively.

In comparison with the forward-generalization gradient, the backward gradient was slightly elevated in the middle. An analysis of variance based on the data of the two middle points of the dimensions showed that forward generalization did not differ from backward generalization [for F_1, F_2 versus B_1, B_2, $F(1, 13) < 1$]. However, the performance difference between the two middle points of the dimensions was significant [for F_1, B_1 versus F_2, B_2, $F(1, 13) = 7.36, p < .05$], indicating that the gradients were fairly sharp.

The question of *program* similarity refers to the manipulation for group FP. The lower part of Table 2 summarizes the forward- and backward-generalization gradients obtained when the standard program per se was systematically varied in the test trials (again, by methods *a* through *e* above). In both cases, the time required to complete a test program increased as the test differed gradually from the standard, resulting in typical, smooth gradients [$F(3, 39) = 7.74, p < .001$; and $F(3, 39) = 9.33, p < .001$] for both forward and backward generalization respectively.

By restricting the data to the two middle points of the dimensions, an analysis of variance revealed that the forward generalization did not differ from the backward generalization [for F_1, F_2 versus B_1, B_2, $F(1, 13) = 3.32$] and that the two middle points of the dimensions did not produce a performance difference [for F_1, B_1 versus F_2, B_2, $F(1, 13) = 1.81$].

In brief, the generalization gradients obtained along the dimension of path similarity were steeper than those obtained along the dimension of program similarity. As is evident from Table 2, the two types of gradients crossed each other. It took group FPS more than four times as many seconds to complete an entirely different program than it did to complete

Table 2. Forward- and backward-generalization gradients for different types of similarity

		Standard	F_1 or B_1	F_2 or B_2	Different
Path similarity (Group FPS)					
Forward generalization	*M*	3.99	12.22	14.83	16.51
	SD	.95	3.75	4.01	4.26
Backward generalization	*M*	3.99	13.26	15.14	16.51
	SD	.95	3.29	4.75	4.26
Program similarity (Group FP)					
Forward generalization	*M*	8.23	12.26	12.33	13.60
	SD	2.53	3.63	3.81	4.05
Backward generalization	*M*	8.23	10.70	11.17	13.60
	SD	2.53	2.58	2.39	4.05

Note: F_1 or B_1 refers to a program whose last or initial direction differs from the standard; F_2 or B_2 to a program whose last or initial two directions differ from the standard. Data in seconds of time to complete the respective programs.

the standard program, whereas it took group FP less than twice as many seconds to complete an entirely different program than it did to complete the standard program.

The finding of the steeper generalization gradients along the dimension of path similarity is easily understood, since the experimental manipulation of path similarity actually involved both path and program. That is, the program itself was manipulated in order to induce changes in the path it defined, whereas the experimental manipulation along the dimension of program similarity involved only the program per se inasmuch as the path varied for the standard program as well as for any test program. Therefore, the difference between two points on the dimension of path similarity (group FPS) was actually much larger than the difference between the two corresponding points on the dimension of program similarity (group FP). It naturally follows that the larger differences between the standard and test programs for group FPS produced greater decreases in performance level.

The finding that—with time required to complete the standard program as base level—it took much longer to complete an entirely different program in group FPS than in group FP is also consistent with the findings of studies on problem-solving set (Gardner and Runquist, 1958). In those studies, the subjects were given a series of problems to be solved successively by a given mode of attack and then tested on a problem for which the mode was inappropriate. In a similar sense, the subjects of group FPS used a much more fixed mode of attack during the training trials.

Finally, the finding of no significant difference between forward and backward generalization implies that changes introduced at the beginning of the program for a chain of instrumental acts behave in approximately the same way as changes introduced at the end. However, given the relatively small number of subjects used in the present experiment, it can be safely concluded only that the difference may not be large.

CONCLUSION

The task used in the present study is hardly classifiable as one of skill learning, since the motor skill involved is minimal. It belongs, rather, to the domain of problem solving. Still, the subject knows exactly what to do for each direction in a program. It is indeed a very simple problem-solving task, although the complexity of a program may be increased without limit. Mowrer (1950) aptly classified instrumental learning as problem solving. According to the present results, some problem-solving variables such as set behave quite similarly.

Because the chain of acts demanded by the sequence of directions in a program leads the subject to a goal, the task used in the present study properly represents a kind of instrumental learning. From the standpoint of experimental analysis, the present study separated two aspects of this learning: that of the space within which a chain of instrumental acts takes place, and that of the program which dictates the nature of those acts. Experiments I and II studied the properties of the space. Both the size of (the information in) the space and the locus of the path of acts in that space affect learning. Experiments III and IV studied the characteristics of the program. Both a program per se and a program with a fixed path define chains of instrumental acts, and the generalization gradients determined by these two kinds of chains have different slopes.

Notes

This study was supported by a grant from the National Science Council to the first author. Experiments I and II were conducted by the second author in partial fulfillment of the requirements for the M.S. degree under the direction of the first author. The assistance of Sin-tsi Lin and Mei-chih Li in collecting the data of Experiments III and IV respectively is gratefully acknowledged. Received for publication December 11, 1970.

References

Ammons, C. H. 1955. Task for the study of perceptual learning and performance variables. *Perceptual and Motor Skills* 5:11–14.

Bakan, P. 1955. Effect of set and work speed on time estimation. *Perceptual and Motor Skills* 5:147–148.

Gardner, R. A., and Runquist, W. N. 1958. Acquisition and extinction of problem-solving set. *Journal of Experimental Psychology* 55:274–277.

Mowrer, O. H. 1950. *Learning theory and personality dynamics.* New York: Ronald Press.

American Journal of Psychology
1971, Vol. 84, No. 4

DIMENSIONAL SHIFTS WITH VERBAL UNITS

Gale Bishop and Jerome Smith
University of Connecticut

College-student subjects trained on a CVC discrimination problem with a unidimensional solution were next given a transfer problem solvable by intra- or extradimensional shift. They then sorted decks of CVCs along the four experimentally manipulated dimensions: meaningfulness, rhyming similarity, color, and size. The intradimensional shifts were easier but produced increasingly longer sorting times, and these transfer effects showed up on all four dimensions.

Contemporary theories of visual-discrimination learning have shifted away from a concern with the response to specific positive and negative cues in a discriminative situation (Hull, 1943; Spence, 1936) and have moved toward an emphasis on the mediational response to common stimulus dimensions of both cues (Kendler and Kendler, 1962; Lovejoy, 1969; Zeaman and House, 1963). The mediational approach recognizes the role of perceptual dimensions in discrimination learning and forms a useful explanatory base in animal (Mackintosh, 1965), child (Shepp and Turrisi, 1966) and adult (Isaacs and Duncan, 1962) studies. It has proven useful with different subject populations, with different kinds of visual materials (Zeaman and House, 1963), and with other sense modalities (Brown and Smith, 1967; Smith, Anderson, Cunningham, and Sjoberg, 1967).

But the mediational view has not found so congenial a home in the area of discrimination of verbal units. Bogartz (1965) was able to demonstrate differences between reversal and partial-reversal shifts when the CVCs used as stimuli lacked the obvious dimensional differences that are comfortably handled by a mediation theory. Kendler, Kendler, and Sanders (1967) found the expected differences between reversal and half-reversal shifts when words were the stimuli but could only find a suggestion of predicted effects when CVC trigrams were the stimuli. Indeed, the verbal-discrimination task itself (Ekstrand, Wallace, and Underwood, 1966) is designed so that perceptual dimensions other than frequency of item occurrence cannot form a basis for discrimination. Ekstrand et al. do not deny that other perceptual dimensions may serve as a basis for discrimination of verbal units, but their frequency theory was formulated to account for performance in their task.

The present study was designed to discover whether the effects predicted by a mediational view and found in a diversity of situations are also to be found when verbal units are discriminated along dimensions other than frequency. The dimensions under investigation were of two types, those naturally associated with verbal stimuli (the 'linguistic' dimensions of meaningfulness and rhyming similarity) and certain 'nonlinguistic' dimensions (color and size).

The role of mediation was studied by means of the intradimensional-shift and extradimensional-shift paradigms (Zeaman and House, 1963). In the intradimensional shift the dimension relevant to the solution of a problem in acquisition remains relevant to the solution of a problem in transfer, although different positive and negative cues are employed. In the extradimensional shift a dimension that was irrelevant during acquisition is made relevant in transfer. Since extradimensional shifts require a change of attention from one perceptual dimension to another, they should prove more difficult than intradimensional shifts and, but for some special circumstances, they do (Shepp and Turrisi, 1966). Dimensional potency was examined by looking at ease of solution of the original CVC problems for the four different dimensions. Further tests of dimensional effects were made by asking for a series of 'sorts' of new items until all dimensions systematically varied had been discovered by the subject.

METHOD

—Subjects—The subjects were 240 male and female students in introductory psychology classes. Participation in the experiment partially fulfilled the course requirements.

–Materials—Each of 16 CVC trigrams was block printed and centered on an unlined 4-by-6-in. index card. A stimulus card was assigned one of two possible values along each of four stimulus dimensions. The dimensions employed were: the color of the ink in which the CVC was printed, the size of the letters, meaningfulness as defined by Noble's m' (1961), and rhyming similarity to the other items in the card deck. The two values along each dimension were: gray or blue, 1¼- or 2¼-in.-tall letters, high meaningfulness (m' from 3.74 to 4.22) or low meaningfulness (m' from 1.04 to 2.50), and high similarity (items ending in AD) or low similarity (items ending in vowel-consonant combinations other than AD and having no consistent similar elements). Two additional decks of 16 cards each were prepared with the same dimensional variations but with new trigrams. The same high and low m' levels were chosen, with the highly similar trigrams ending in AY for one deck and UN for the other. Color values were orange or purple in one case and yellow or green in the other. Size differences were 1½ or 2½ in. and 1¾ or 2¾ in.

—Procedure—During *training* each subject was randomly assigned to one of the four stimulus dimensions with the restriction that each group contain equal *N*s. A deck of cards was selected and presented, one card at a time, with instructions that the subject classify the word as an 'X' or a 'Y' word according to a rule the subject himself was to discover. Knowledge of results was given after each card was presented, and a trial was defined as a single run through the deck of 16 cards. The decks used were counterbalanced so that all three decks were equally represented in each of the four groups. Cards were shuffled after each trial, and the training criterion was set at one errorless trial. After reaching criterion the subject was given one additional, overlearning, trial and requested to verbalize his solution. If criterion was not reached in five trials, he was told the appropriate dimension for categorization and given one more trial.

During transfer ('*shift*'), each subject was given a new deck, with instructions identical to those during training. For half the subjects the appropriate category remained the same as it had been (intradimensional shift). For the other half, one of the three remaining dimensions was chosen (extradimensional shift) so that each of the new dimensions was equally represented. Again criterion was set at one errorless trial and the correct category was provided by the experimenter if criterion was not reached within five trials.

After the transfer task came another task ('*sort*'), in which each subject was given a third deck of 16 cards and asked to sort it into two equal piles according to any organizational rule he chose ('sort number 1'). When he finished he was asked to re-sort the same deck on the basis of a different rule (sort number 2). A maximum of six sorting opportunities (up through sort number 6) were given to arrive at the four sorts representing the dimensions of color, size, meaningfulness, and rhyming similarity. Time measures were taken from the time the subject got the deck until he completed a sort. If during any sorting period he went for 3 min without finishing a sort, he got a hint: the experimenter chose two cards distinguishable along all the dimensions on which the subject had not as yet sorted and told him that they belonged to two categories. If the subject could not then sort for 2 min after the hint, he was given increasingly obvious clues and questioned until he sorted the deck in a fashion other than he had previously. The time required for hints was included as part of the particular sorting time.

RESULTS

Data for the training and shift portions of the experiment are presented in Table 1. Both mean trials to criterion and mean errors to criterion are included. Since the two measures were highly correlated and analyses done on them produced similar outcomes, only the trials-to-criterion analyses are reported.

Table 1. Mean trials (and errors) to criterion

	Training		Shift	
	Intra-dimensional	Extra-dimensional	Intra-dimensional	Extra-dimensional
Color	4.5 (22)	4.4 (21)	1.8 (1)	3.7 (19)
Size	5.3 (31)	5.5 (31)	1.63 (1)	4.8 (28)
Rhyming similarity	3.03 (8)	3.1 (9)	2.36 (5)	3.6 (16)
Meaningfulness	4.46 (22)	4.2 (22)	2.16 (4)	4.0 (21)

Training

A 4 (dimensions) × 2 (shift conditions) factorial analysis of variance revealed significant dimensional effects $[F(3, 232) = 21.63, p < .01]$. Since the designations of intra- and extradimensional here refer only to the fact that the subjects were treated differently during the *shift* portion of the experiment, no significant effects of shift or of the interaction of shift and dimension were expected, nor were they found (both Fs $<$ 1). The dimensional effect was further analyzed by means of Scheffe comparisons, which revealed that rhyming similarity was most easily learned ($F > F'$ at $\alpha = .05$), size was hardest to learn ($F > F'$ at $\alpha = .05$), and that size and color combined were more difficult than meaningfulness and rhyming similarity combined ($F > F'$ at $\alpha = .05$).

Shift

In the shift portion of the experiment, intra-extradimensional differences were expected, and a 4 (dimensions) × 2 (shift conditions) analysis substantiated the predicted ease of the intra- as opposed to the extradimensional condition $[F(1, 232) = 121.8, p < .01]$. The dimensional effect that appeared in the training portion of the study was washed out $[F(3, 232) = 1.22, p > .05]$ by the interaction of dimension and shift condition $[F(3, 232 = 4.83, p < .01]$. The interaction is clearly illustrated by order of ease of problem solution in the shift data of Table 1. Intradimensional subjects found size easiest, followed by color, m', and rhyming similarity, while extradimensional subjects reflected the training preferences and found rhyming similarity easiest, followed by color, m', and size.

Sort

The time to sort for each shift condition on each of the first four sorts is presented in Figure 1. Intra- and extradimensional groups both required increased sorting time as sort number increased, but the intradimensional groups required increasingly more time. A 2 (shift) × 4 (sort number) analysis of variance produced a significant effect of shift $[F(1, 238) = 20.3, p < .01]$, a significant effect of sort number $[F(3, 714) = 104.2, p < .01]$, and a significant interaction of shift condition and sort number $[F(3, 714) = 3.7, p < .05]$. Although the particular dimension sorted is confounded with sort number in this analysis, the rationale for collapsing across dimensions and analyzing shift × sort-number data was provided by an unsystematic distribution of dimensional preferences across the four sorts. Further support for the greater sorting difficulty of the intradimensional subjects is given by the number of hints they needed (78, compared to 48 for the extradimensional subjects).

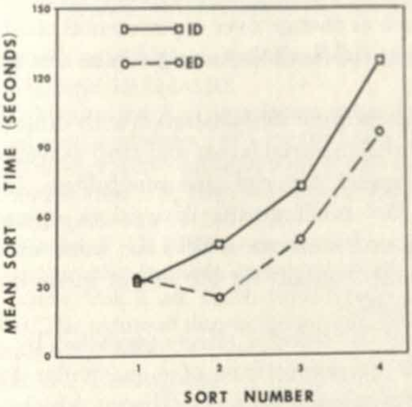


Fig. 1. Sorting time for intradimensional (ID) and extradimensional (ED) groups; data collapsed across dimensions

DISCUSSION

The experimental findings (Table 1) clearly support a mediational approach to discriminations involving verbal units. Despite strong dimensional preferences during training, changes in dimensional relevance (extradimensional shifts) were always more difficult than changes only in the cues of the dimension relevant during training (intradimensional shifts). Further indication of a broad kind of mediation is provided by the time-to-sort data in Figure 1. The intradimensional subjects had greater difficulty finding new dimensions to sort than did the extradimensional subjects, a difficulty attested to by their increasingly longer sorting times and their need for more hints. This difference may be thought of as a kind of learning-set difference, or as attention to a complex dimension such as change over experimental conditions, or perhaps in terms of experimental flexibility or rigidity as discussed by Schroeder and Rotter (1952).

Meditational effects were demonstrated with dimensions not usually associated with verbal material (color and size) as well as with attributes more natural to verbal material (meaningfulness and rhyming similarity). The fact that nonlinguistic dimensions operated in the same fashion as linguistic dimensions within the same experimental framework provides further support for the role of mediation in the discrimination of verbal units.

Demonstration of the transfer effects predicted by mediation theory does not establish the importance of a particular type of mediation theory, nor does it negate the frequency theory, which was formulated to account for the verbal-discrimination situation in which only frequency affords a basis for discrimination. However, it would seem reasonable to say that finding the same kind of intra-extradimensional difference with verbal units as those found with other discriminanda speaks to the generality of theories that predict such transfer effects.

Notes

Reprint requests may be sent to Jerome Smith, Department of Psychology, U-20, University of Connecticut, Storrs, Connecticut 06268. Received for publication January 29, 1971.

References

Bogartz, W. 1965. Effects of reversal and nonreversal shifts with CVC stimuli. *Journal of Verbal Learning and Verbal Behavior* 4:484–488.

Brown, Z., and Smith, J. 1967. Tactual-kinesthetic discrimination shifts in retarded children. *Journal of Experimental Child Psychology* 5:455–462.

Ekstrand, B. R., Wallace, W. W., and Underwood, B. J. 1966. A frequency theory of verbal discrimination learning. *Psychological Review* 73:566—578.

Hull, C. L. 1943. *Principles of behavior*. New York: Appleton-Century-Crofts.

Isaacs, I. D., and Duncan, C. P. 1962. Reversal and nonreversal shifts within and between dimensions in concept formation. *Journal of Experimental Psychology* 64:580–585.

Kendler, H. H., and Kendler, T. S. 1962. Vertical and horizontal processes in problem solving. *Psychological Review* 69:1–16.

Kendler, H. H., Kendler, T. S., and Sanders, J. 1967. Reversal and partial reversal shifts with verbal material. *Journal of Verbal Learning and Verbal Behavior* 6:117–127.

Lovejoy, E. 1968. *Attention in discrimination learning.* San Francisco: Holden-Day.

Mackintosh, N. J. 1965. Selective attention in animal discrimination learning. *Psychological Bulletin* 64:124–150.

Noble, C. E. 1961. Measurement of association value (a), rated association (a') and scaled meaningfulness (m') for the 2100 CVC combinations of the English alphabet. *Psychological Reports* 8:487–521.

Schroeder, H. M., and Rotter, J. B. 1952. Rigidity as learned behavior. *Journal of Experimental Psychology* 44:141–150.

Shepp, B. E., and Turrisi, F. D. 1966. Learning and transfer of mediating responses in discriminative learning. In *International review of research in mental retardation,* Vol. 2, ed. N. R. Ellis. New York: Academic Press.

Spence, K. W. 1936. The nature of discrimination learning in animals. *Psychological Review* 43:427–449.

Smith, J., Anderson, D., Cunningham, T., and Sjoberg, W. 1967. A comparison of auditory and visual discrimination learning in retardates. *American Journal of Mental Deficiency* 72:445–449.

Zeaman, D., and House, B. J. 1963. The role of attention in retardate discrimination learning. In *Handbook of mental deficiency,* ed. N. R. Ellis. New York: McGraw-Hill.

ANNOUNCEMENT

Senior Fulbright-Hays Program for 1973–1974

The Committee on International Exchange of Persons announces that applications for senior Fulbright-Hays awards for lecturing and research during 1973–1974 in over 75 foreign countries will be accepted in the spring of 1972. Specialists in the field of psychology who are U.S. citizens and have a doctorate or college teaching experience are invited to indicate their interest in these 1973–1974 awards by immediately completing a simple registration form, available on request from

Senior Fulbright-Hays Program
2101 Constitution Avenue
Washington, D.C. 20418

Registrants will receive the detailed announcement of available awards as soon as it is issued in the spring, in time to consider the possibilities and to apply before the closing date. July 1, 1972, is the deadline for applying for research awards, and it is the suggested date for filing for lectureships.

Senior Fulbright-Hays awards generally consist of a maintenance allowance in local currency to cover normal living costs of the grantee and family while in residence abroad, and round-trip travel for the grantee (transportation is *not* provided for dependents). For lecturers going to most non-European countries, the award includes a dollar supplement, subject to the availability of funds, or carries a stipend in dollars and foreign currency, the amount depending on the assignment, the lecturer's qualifications, salary, and other factors.

BOOK REVIEWS

Julian Hochberg
Columbia University

The Mechanisms of Perception

By Jean Piaget. New York: Basic Books, 1969. Pp. xxix, 384. $12.50.

This translation of *Mecanismes perceptifs* covers some 20 years of Piaget's work in the area of perception, most of which originally appeared under the heading of *Recherches sur le développement des perceptions* in the *Archives de Psychologie*, between the years of 1942 and 1961. It is an attempt to explain a large body of data with one general theoretical system. In order to accomplish this task, given the broad range of topics discussed, Piaget is forced to leave many of the central notions of his theoretical system rather loosely and incompletely defined. He does make an attempt to specify his theoretical position more precisely by using a system of algebraic statements that he refers to as a "relational language," but his symbols and relationships suffer from the same lack of specificity as his verbal constructs. In addition, the algebraic statements tend occasionally to obscure simple relationships that could be more clearly stated in words. Despite these difficulties, *The Mechanisms of Perception* has many interesting and exciting elements.

The heart of Piaget's perceptual system is an active observer who is constantly exploring the visual environment. The act of fixating a target, or shifting one's attention to it, is called *centration*. Centration results in a systematic sampling of the elements of the visual array. When an element is sampled, it is called an *encounter*. When the eye is static, elements are successively sampled, and this sampling may be facilitated by microsaccades. *Transports*, or active eye movements, also result in encounters. It is postulated that the number of encounters determines the perceived length of a line. More encounters tend to occur in the region of the field that is centered; this results in an overestimation of the size of the fixated target. In psychophysical experiments where there is a standard stimulus, the additional centrations result in the 'error of the standard,' or overestimation of the standard relative to comparison stimuli.

The observer is not only actively moving his eyes, but he is actively making comparisons between elements in different parts of the array. These comparisons, which are called *couplings*, result in perceptual over- and underestimations, which operate in the same direction as size contrast when they are incomplete. If all possible comparisons were made between two adjacent elements, it is hypothesized, this type of perceptual distortion would be greatly reduced.

Since Piaget is dealing with relative over- and underestimations of the size of targets in the field, it is only natural that the primary focus of his theorizing and research is on the classical geometric illusions. He divides illusions into two classes, those which require an active observer for their occurrence and those which do not. He reasons that if an illusion figure is briefly presented in a tachistoscope, there is probably only enough time for a single centration to occur. If the illusion occurs under these conditions, it is said to be due to *field effects*, which are simply the relational properties of the visual array. These are called *primary illusions*. Unfortunately, some of the relational aspects required for theoretical coherence are not the actual lines in a figure, but virtual lines and constructions that must be mentally provided by the observer. The comparison of real lines to imaginary geometrical constructs is a major weakness in the interpretation of the primary illusions. *Secondary illusions* require active exploration and comparison. The hypothesis here is that these illusions increase in strength with increased exposure duration, since more time is then available for comparison and interaction.

The distinction between primary and secondary illusions provides the basis for Piaget's developmental thrust. Since secondary illusions depend upon exploratory activity and since this activity becomes more important and systematic with age, there is a developmental increase in the magnitude of the secondary illusions. These same exploratory activities increase the likelihood of a more complete comparison of all relative attributes of a figure and hence reduce the magnitude of the primary illusions; thus, primary illusions decrease with age. The hypothesized difference in exploratory activity as a function of age is supported with eye-movement data. Piaget also mentions an interesting class of illusions in which the developmental trend rises to a maximum and then declines. The argument here is that the younger observers lack the ability to achieve closure in complex figures and hence the relational attributes available to older observers are not yet apparent to them. After the ability to effect closure has been achieved, the exploratory mechanisms begin to compensate for the presence of the primary illusion. Data are provided to show that children have difficulty in achieving closure.

At this juncture Piaget introduces another major theoretical construct, the notion of a *scheme*. A scheme is simply a generalization, set, expectation, or organizational predisposition that may arise as a function of repeated exposure to certain stimulus relationships. Thus, the Gestalt 'good figures' are seen as perceptual schemes. These schemes may now also serve to alter the nature of the percept. For instance, incorporating the Müller-Lyer illusion as part of a square reduces the magnitude of the distortion, since the scheme associated with squareness competes with the illusion. And with this arsenal of theoretical constructs, Piaget now extends the range of his theorizing to include the perceptual constancies, causality, and the perception of motion, velocity and time—all of which, he contends, result from the construction of perceptual schemes based upon exploratory activity.

The last 86 pages of the book are an attempt to show that perception and intelligence are related only in the sense that they both depend upon an active exploring observer, capable of assimilating information into conceptual as well as perceptual schemes. This discussion is mainly theoretical, and the experiments reported are only vaguely relevant to the issue.

The major contribution of this volume, in addition to bringing Piaget's perceptual work to the attention of the English-speaking community, is that it summarizes a massive amount of research on visual illusions, much of which contains developmental comparisons. In addition, its basic theoretical tenet—that the organism is continually active, constantly exploring the environment with eye movements, mental comparisons, and overt motor movements—provides an interesting restatement of the basic empiricist-transactional viewpoint. The major weaknesses of the volume lie in the looseness or absence of definition of some of the major constructs, the obscurity of some of the algebraic manipulations, and the failure to discuss possible alternative explanations in the course of his presentation. Most workers in the field will probably find that the positive aspects of this contribution far outweigh the negative.

Stanley Coren, *New School for Social Research*

The Perception of People and Events

By Peter B. Warr and Christopher Knapper. New York: Wiley, 1968. Pp. xii, 445. $10.00.

The term 'perception' is used very broadly here, so that it shades off into impression formation, attribution, trait aggregation, inference, and implicit personality theory. Nevertheless, Warr and Knapper maintain throughout that perception of people and of objects will ultimately be explained by the same higher-order principles, and they devote an entire chapter to similarities between object constancy and constancy in "person perception." Such comparisons are facilitated by a general information-processing model, which also provides an organizational format for the book. Within this format two very different projects are undertaken simultaneously: to report scores of original experiments conducted during a three-year investigation of factors related to British newspaper accuracy, and to meet the need for a comprehensive literature review and textbook on person perception.

Previous research in this area can be divided roughly into two categories: studies of perceptual accuracy and studies of perceptual processes. Like most contemporary researchers, Warr and Knapper have focused on the latter, emphasizing the following sequential components: (*a*) input, in the form of present stimulus information, present context information, and "stored stimulus person information"; (*b*) an "input selector," which is affected by a perceiver's stable characteristics and current state: (*c*) a processing center, which contains inference rules and combination rules and corresponds to what other researchers have called implicit personality theories; (*d*) output, in the form of affective and attributive responses and expectancies, which in turn serve as inputs to subsequent perception, to long-term memory, and to behavior formulation. This model, for which an elementary computer-simulation program has been written, aids substantially in pulling together what have often seemed isolated research topics, thus making it more apparent where further study is needed.

In discussing the attributive component of their model, Warr and Knapper distinguish between *episodic* judgments (regarding someone's state during a particular sequence or episode of behavior) and *dispositional* judgments (regarding someone's chronic state), which latter may refer to either overt or covert attributes (say, 'fingernail biting' versus 'being a nervous person'). Most studies of person perception have focused on dispositional judgments of covert characteristics, perhaps because these are crucial in making clinical assessments and are assumed central to other forms of social interaction as well. The authors' research fits comfortably into this tradition. In a typical study they asked subjects to read about a person—often a British political figure such as Harold Wilson—performing in a particular situation—such as delivering a speech—and then to give their impressions of him on several semantic-differential scales. Typically, these included a few scales loading highly on each of the well-known evaluation, potency, and activity factors, plus unique scales relevant to the particular hypothesis under investigation. This paradigm was then altered in numerous ways to reveal effects of prior conceptions (that "stored stimulus person information"), of sex, personality, religion, and so on. In all, some 40 hypotheses were tested.

Because this research grew primarily out of questions about newspaper accuracy, most of it involved *communication* about a social stimulus rather than direct encounter; no social *interaction* with the stimulus person was involved. As a result, the relationships between person perception and social behavior, as well as currently controversial hypotheses about self-perception and attribution of emotional states, receive less attention here than they deserve. This is not meant as a criticism of the authors' research program. In a modern industrialized society much of one's information about socially significant events comes to him indirectly, already filtered, amplified, edited—already selectively perceived, sometimes by friends, often by professionals in the mass media. (The authors cite a United Nations study indicating that most people in the Western world consult at least one newspaper each day.) Thus it is extremely important to know how perception of public figures, and of their behavior in specific instances, is influenced by newspaper articles and photographs, the stimuli most often employed here. But many important questions about social perception and its effects have to do with social interaction, with the behavioral and emotional responses of perceivers to immediately available social stimuli; and these questions are relatively neglected, considering that this book is supposed to serve also as a text. For other reasons as well, the attempt to combine a research monograph with an elementary textbook is unsuccessful. The authors' painstaking, occasionally esoteric experiments are often unsuitable for a novice, as are the attempts to apply symbolic logic to the problem of trait inference and as are the preliminary efforts to write a computer-simulation program. The choice of newspaper articles as the main communication medium limits the range of issues considered and at the same time raises other issues (such as the influence of headlines and typography) that will not interest most readers.

A second weakness of the book stems from the authors' heavy reliance on semantic-differential ratings. Their apt analogy between connotative and denotative meaning, on the one hand, and covert and overt characteristics, on the other, reveals part of the problem. After stressing similarities between the perception of objects and people, and after noting the dearth of studies on the perception of the overt characteristics of people, they proceed to study covert characteristics, using an ambiguous measure. Does a rating on a 'large-small' scale, for example, reflect a judgment about Harold Wilson's physical size or about his stature in some metaphorical sense? Moreover, the generality of the semantic differential makes it unlikely that the specific effects of the proposed model's various components can be disentangled, although it should be noted in the authors' favor that they always analyze results scale by scale rather than from the combined, and ambiguous, factor scores.

Finally it is appropriate to point out that—contrary to expectations raised by the title—there is little in this book about the perception of *events* (with the exception of one football game) and little discussion of how one would assess "event perception." Social psychologists have been relatively careless about conceptualizing complex social organizations, social events (such as mass demonstrations), and physical events (such as earthquakes) simply as 'stimuli.' Unfortunately, Warr and Knapper are not exceptional in this regard.

Despite such limitations, this is an important and useful book, especially for researchers. The extended discussion of semantic-differential technique is a worthwhile supplement to the available literature, particularly for students of social perception. The choice of journalism as a source of stimulus materials adds unusual 'ecological' validity to some of the authors' studies. The literature review and the integration of previous findings make a significant contribution to current efforts to understand person perception, attribution, and interpersonal communication—topics which may soon displace attitudes and attitude change as the central concern of experimental social psychology.

Philip Shaver, *Columbia University*

Studies in Animal and Human Behavior, Vol. 1

By Konrad Lorenz. Translated by Robert Martin. Cambridge, Mass.: Harvard University Press, 1970.

This is the first volume of Konrad Lorenz's early papers written during the years 1931–1942. They are excellently translated by Robert Martin, himself an ethologist, and credit goes to him and to the fact that he was able to discuss questions pertaining to terminology with Lorenz himself. The book also contains almost 10 pages of Lorenz's later notes and comments on his own papers, thus showing the reader how he regards many of his earlier ideas today.

Lorenz has written an extensive introduction in which he comments not only on the evolution of his ideas but on how the basic ideas of ethology have been received and on how much of the criticism, centered around his own person, is characterized by a passion not usually found among scientists who disagree among themselves. Lorenz has used this opportunity to reply to his critics and has also commented on the more far-out speculations by Desmond Morris in *The Naked Ape* and by Robert Ardrey in his books—authors who, Lorenz says, are causing him serious alarm. He talks about his early development as an ethologist and about the influence that O. Heinroth, C. O. Whitman, and Wallace Craig had on his thinking; and he restates clearly and concisely the basic ethological tenets, which emphasize thorough familiarity with one's animal, the animal one has chosen to study. He pulls no punches, yet he is equally honest in tracing his own groping formulations toward conceptualization, the errors and detours that were a part of the growth of the modern ethology he so profoundly influenced and shaped.

The first paper, "Contributions to the Study of the Ethology of Social Corvidae" (1931) is, in Lorenz's own words, "a good illustration of how that kind of knowledge [of the animal one is studying] is gained." He calls the paper "a plain story, pure description, with precious little theorizing." He did write about species-specific drive activity there, though, and made his first mention of imprinting and of how differently it is expressed in jackdaws as opposed to magpies. In this paper we see his acceptance at that time of classical reflex theory as stated by Sherrington, and he relates how it took him six years to realize that his own observations were incompatible with Sherringtonian reflexology and with Ziegler's theory that innate motor patterns were chain reflexes.

In 1932 "A Consideration of Methods of Identification of Species-specific Instinctive Behavior Patterns in Birds" was published. It is an account of the characteristics he employed in recognizing instinctive behavior patterns. It was written, according to Lorenz, at a time when he was not familiar with the writings of purposive and behavioristic psychologists in the United States.

The third paper in this book is the famous 'companion' paper. Published in 1935, it was entitled "Companions as Factors in the Bird's Environment" and was one of the first of his papers to be translated into English. It is primarily regarded as the paper in which Lorenz stated the basic criteria of the phenomenon of imprinting, a paper which can justifiably be said to have stimulated massive research in Europe and the United States from the 1950s on. The paper contains other important concepts—the concepts of the parent, infant, sexual, social, and sibling companion, of the innate schema, releaser, function, and so on. There is a detailed discussion of the instinct problem, leading into a discussion of McDougall's ideas. These latter were to be treated in more detail in his paper "The Establishment of the Instinct Concept" in 1937, as were Spencer, Lloyd Morgan, and the theory that "instinct, by becoming more and more strongly influenced by learning and intelligence, merges into the latter ontogenetically as well as phylogenetically." Lorenz also used the idea of the spontaneity of behavior in his critique of the chain-reflex theory of behavior and dismissed the 'instinct theory' of the behaviorists because of their lack of detailed knowledge of their animals.

The fourth paper, "Taxis and Instinctive Behavior Pattern in Egg-Rolling by the Greylag Goose," appeared in 1938. Here we find the first clear formulation of the orienting responses (or taxes, as they came to be called), which are reflex in nature, and of the innate motor pattern, which is physiologically something quite different. Lorenz credits his correspondence with Wallace Craig and the work of the physiologist E. V. Holst for his clear conceptualization of the ideas of appetitive behavior, innate releasing mechanism, and innate motor patterns, ideas which even today are such important concepts in ethology. 'Egg-rolling' to retrieve an egg outside the nest is analyzed in detail, and various releaser characteristics, the habituation of the response, and so forth are discussed.

The final paper in this book, "Inductive and Teleological Psychology," was published in 1942. It is a response to Bierens de Haan's critique of Lorenz's theories. De Haan espoused all-encompassing explanatory concepts like 'instinct,' and we see Lorenz contrasting these teleological 'holistic considerations' with inductive scientific research.

Those who have accused Lorenz over the years of labeling some behaviors as instinctive and accepting that as adequate explanation which requires no further research would do well to read this last-named paper. Those who are familiar with ethological work today will find that the evolution of the ideas traced in all these papers provides a ringside seat, as it were, of the work of Konrad Lorenz. And for all those who are interested in ethology and its relevance to other areas of behavioral study, this book should be required reading, since so many of the questions asked about ethology can be understood by reading these pages. I hope that the second volume will be available soon.

Erich Klinghammer, *Purdue University*

Verbal Conditioning and Behavior

By J. P. Das. New York: Pergamon Press, 1969. Pp. vii, 163. $7.00.

By the author's own acknowledgment, this book is not about what the title would suggest; namely, verbal operant conditioning related in some degree to Skinner's work. Skinner's name cannot even be found in the author index, although "Skinnerian" appears a few times in the text. A Hull-Spence orientation is claimed but freely departed from, and the statement "we are eclectics" (p. 15) seems to fit the case.

The book consists essentially of a summary of more than 20 research papers published during the previous 10 years by the author singly or with co-authors. At least one study has been done in each of a number of established research areas bearing in one way or another on the "problem of meaning," which, Das concludes, is "the central theme of this book" (p. 134). Topics treated include classical and instrumental conditioning, awareness, probability learning, mediation and transfer of meaning, semantic satiation, hypnosis as verbal conditioning, and neurodynamics of verbal conditioning. Chapters are not tied together, and there is no clear line of development, except that there is an increase in the level of complexity of subject matters as the book progresses. Most of the papers were published in journals readily available to those likely to be interested. They were found, upon sampling, better organized and more readable than the warmed-over versions presented in this book.

Ralph F. Hefferline, *Columbia University*

The Phenomenology of Husserl: Selected Critical Readings

Edited and translated by R. O. Elyeton. Chicago: Quadrangle Books, 1970. Pp. 306. $15.00; paperback, $2.95.

This volume presents six essays primarily concerned with the development of Husserl's philosophy from the turn of the century to his death in 1938. Although the papers are dated from 1930 to 1962, none had appeared in translation until this volume. The contributors to the volume are either former students of Husserl (such as Oskar Becker and Eugene Fink) or philosophers who have been involved in the preparation of his posthumously published collected works. The essays are all concerned with tracing the development of transcendental phenomenology.

On the whole, this volume will be of little interest to psychologists, since the relationship of phenomenology to psychology is barely mentioned and the major concern is with metaphysical issues. Of the six essays, the ones by Oskar Becker and Hans Wagner come closest to being relevant, as they contain an excellent presentation of phenomenological reduction—Husserl's major methodological innovation and the key to his system.

Walter Cohen, *State University of New York at Buffalo*

Books received

(The books listed here have not as yet been noted in our pages. Their listing does not preclude their later review, however.)

Allen, David. *The Price of Women.* New York, N.Y.: Jarrow Press, 1971. Pp. 206. $6.95.

Anglin, Jeremy M. *The Growth of Word Meaning.* Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xiii, 108.

Arneheim, Rudolf. *Entropy and Art: An Essay on Disorder and Order.* Berkeley and Los Angeles: University of California Press, 1971. Pp. 64. $4.50.

Ball, John C., and Chambers, Carl D., eds. *The Epidemiology of Opiate Addiction in the United States.* Springfield, Ill.: Charles C Thomas, 1971. Pp. xvi, 327. $15.50.

Bar-Adon, Aaron, and Leopole, Werner F., eds. *Child Language: A Book of Readings.* Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xvi, 477. $11.95.

Baughman, E. Earl. *Black Americans: A Psychological Analysis.* New York: Academic Press, 1971. Pp. xxi, 113. $6.95.

Belenky, Robert. *Fragments of a Lesson Plan.* Boston: Beacon Press, 1971. Pp. x, 214. $7.50.

Bergel, Franz, and Davies, D. R. A. *All about Drugs.* New York: Barnes & Noble, 1971. Pp. x, 206. Paperback, $2.50.

Borger, Robert, and Cioffi, Frank, eds. *Explanation of the Behavioral Sciences.* New York: Cambridge University Press, 1971. Pp. xii, 520. $15.00.

Brown, George Isaac. *Human Teaching for Human Learning: Introduction to Confluent Education.* New York: Viking Press, 1971. Pp. xvii, 298. $8.50.

Brozek, Joseph, and Sibinga Maarten, S., eds. *Origins of Psychometry: Johan Jacob de Jaager on Reaction Time and Mental Processes.* Nieuwkoop, Netherlands: DeGraaf, 1970.

Brussel, James A., and Irwin, Theodore. *Instant Shrink: How to Become an Expert Psychiatrist in Ten Easy Lessons.* New York: Cowles, 1971. Pp. xv, 136. $4.95.

Cartwright, Desmond S., and Cartwright, Carol I. *Psychological Adjustment: Behavior and the Inner World.* Chicago: Rand McNally, 1971. Pp. xi, 248. Paperback, $4.95.

Chaundler, Christine. *Every Man's Book of Superstitions.* New York: Philosophical Library, 1971. Pp. 174. $6.00.

Churchill, Don W.; Alpern, Gerald D.; and DeMyer, Marian K., eds. *Infantile Autism*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xviii, 340. $15.50.

Cohen, Joel E. *Casual Groups of Monkeys and Men: Stochastic Models of Elemental Social Systems*. Cambridge, Mass.: Harvard University Press, 1971. Pp. xiii, 175. $8.00.

Cornsweet, Tom N. *Visual Perception*. New York: Academic Press, 1970. Pp. xiii, 474. $15.00.

Curi, Umberto. *Analisi operazionale e operazionalismo*. Padua: University of Padua, 1970. Pp. 235.

Current Trends in Criminological Research. Collected Studies in Criminological Research, Vol. 6. Strasbourg: Council of Europe, 1970. Pp. 495.

DeMott, Donald W. *Toposcopic Studies of Learning*. Springfield, Ill.: Charles C Thomas, 1970. Pp. xi, 262. $13.75.

Dussault, Gilles. *A Theory of Supervision in Teacher Education*. New York: Teachers College Press, 1970. Pp. ix, 275. Paperback, $3.50.

Edgerton, Robert B. *The Cloak of Competence: Stigma in the Lives of the Mentally Retarded*. Berkeley and Los Angeles: University of California Press, 1971. Pp. xv, 233. Paperback, $2.65.

Eisendrath, Craig R. *The Unifying Moment: The Psychological Philosophy of William James and Alfred North Whitehead*. Cambridge, Mass.: Harvard University Press, 1971. Pp. xvi, 290. $10.00.

Elkind, David. *Children and Adolescents: Interpretive Essays on Jean Piaget*. New York: Oxford University Press, 1971. Pp. xi, 160. $5.95.

Fontana, Vincent, J. *The Maltreated Child: The Maltreatment Syndrome in Children*. 2d ed. Springfield, Ill.: Charles C Thomas, 1971. Pp. xv, 96. $7.00.

Friedman, Alfred S.; Sonne, John C.; and Speck, Ross V., eds. *Therapy with Families of Sexually Acting-out Girls*. New York: Springer, 1970. Pp. x, 214. $7.50.

Furrer, Walter L. *Neue Wege zum Unbewussten*. Stuttgart: Hans Huber, 1971. Pp. 285.

Garner, Harry H. *Psychotherapy: Confrontation Problem-solving Technique*. Saint Louis: Warren H. Green, 1970. Pp. xiv, 347.

Gazzaniga, Michael S., and Lovejoy, Elijah P., eds. *Good Reading in Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. ix, 569.

Gergen, Kenneth J. *The Concept of Self*. New York: Holt, Rinehart and Winston, 1971. Pp. x, 110.

Gilgen, Albert R., ed. *Contemporary Scientific Psychology*. New York: Academic Press, 1970. Pp. xiii, 328.

Goguelin, Pierre. *La formation continué des adultes*. Paris: Presses Universitaires de France, 1970. Pp. 196. Fr 13.

Gordon, Edwin, ed. *Experimental Research in the Psychology of Music*. Iowa City: University of Iowa Press, 1970. Pp. vii, 123. $4.75.

Goldenberg, I. Ira. *Build Me a Mountain: Youth, Poverty, and the Creation of New Settings*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xiv, 498. $10.00.

continued on page 472

ANNOUNCEMENT

Summer Institute in Survey Research Techniques

For the twenty-fifth consecutive year, the Survey Research Center of the University of Michigan will hold a Summer Institute in Survey Research Techniques.

The Institute is designed to meet some of the educational and training needs of men and women engaged in business, governmental research and other statistical work, and also to meet the needs of graduate students and university instructors interested in quantitative research in the social sciences.

The 1972 Institute in Survey Research Techniques will be presented in two four-week sessions; the first from July 3 to July 28, and the second from July 31 to August 26. These two sessions may be taken independently or successively.

For further information, please write to

Mrs. Helene Hitchcock
Senior Administrative Assistant
Office of the Director
Survey Research Center
Institute for Social Research
P.O. Box 1248
Ann Arbor, Michigan 48106

INDEX

Rand B. Evans
Wright State University

Authors

(The names of authors of original articles are printed in CAPITALS and SMALL CAPITALS; of authors of books reviewed, in roman; and of reviewers, in *italics*.)

ANDERSON, N. H., 543
Andreas, B. G., 298
APPELLE, S., 487
ARTHUR, B., 355
ATTNEAVE, F., 147
Averill, J. R., 456

BECK, J., 1
Berlin, B., 294
Bertalanffy, L. von, 296
Bever, T. G., 298
BINFORD, J. R., 529
Birren, F., 295
BISHOP, G., 581
Blakemore, C., 455
BOLT, M., 246
Burnham, R. W., 294, 295
Byrne, W., 293

CAIN, W. S., 52
CAREY, S. T., 123
COHEN, J., 23
Cohen, W., 597
COLTHEART, M., 355
Coren, S., 450, 591
COREN, S., 565
Craik, K. J. W., 447
Crovitz, H. F., 141, 142, 449

DAMIANOPOULOS, E. N., 12
Das, J. P., 595
Deese, J., 293
DELPRATO, D. J., 185
Deutsch, J. A., 294
DEY, M. K., 167
DONNELLY, M. B., 407

EFRON, R., 365
Egeth, H., 140
Elkind, D., 142
Elyeton, R. O., 595
EPSTEIN, W., 66
ERIKSON, J. R., 100
ERLEBACHER, A., 477

FISCHER, G. J., 180
Flavell, J. H., 142
FLEMING, J. S., 513
Furth, H. G., 142

GALPER, R. E., 351
Gazzaniga, M. S., 454
George, F. H., 140
Gibson, E. J., 288
Gilinsky, A., 292
GLICK, J., 210
GOTTWALD, R. L., 425
GRUNEBERG, M. M., 473
GUMMERMAN, K., 173

HALL, J. F., 521
HARRIS, L., 218
Hazlitt, H., 141
Hefferline, R. F., 595
HINRICHS, J. V., 268
HOCHBERG, J., 351
Hood, D. C., 452
Howard, I. P., 457
HUI, C., 567
Hunt, E., 447

Jarvinen, J., 452
JOHNSON, G. J., 276
Johnson, H. J., 296

Kay, P., 294
Keppel, G., 293
Kintsch, W., 292
Klinghammer, E., 595
Knapper, C., 591
Kuhn, D., 144

LAACK, W. J., 407

Lee, D. N., 365
Lee, W., 194
Leeuwenberg, E. L. J., 307
Levis, D. J., 555
Liu, I., 567
Loeb, M., 529
London, P., 295
Lorei, T. W., 23
Lorenz, K., 593

McBurney, D. H., 501
McCartney, W., 295
MacKay, D. G., 35
Maier, H. W., 142
Manning, S. K., 421
Martin, E., 123
Mehrabian, A., 145
Millward, R. B., 85
Minsky, M., 445
Moray, N., 452
Moskowitz, H. R., 387
Mostofsky, D. I., 454

Neisser, U., 452

Norman, D. A., 140

O'Connell, R. J., 295
Olson, R. K., 147

Papert, S., 445
Pasnak, R., 235
Pastore, N., 449
Phillips, J. L., Jr., 142
Piaget, J., 589
Postman, L., 293
Postman, L., 461
Prather, D. C., 377
Pribram, K. H., 455, 456

Reber, A. S., 85
Rimoldi, H. J. A., 407
Rosenhan, D., 295
Rothkopf, E. Z., 298
Rubin, M. L., 450

Sanders, A. F., 453
Sasson, R. Y., 253
Schaller, M. J., 218
Sekuler, R., 477
Shaver, P., 593
Sherwood, S. L., 447
Singer, R. D., 145
Singh, D., 134
Slak, S., 205
Smith, J., 581
Stark, K., 461
Steinmetz, J. I., 112
Stevens, J. C., 52
Summers, S. A., 513
Sykes, R. N., 473

Teghtsoonian, M., 75, 437
Teghtsoonian, R., 75, 437
Templeton, W. B., 452
Turnage, T. W., 112

Walls, G. L., 450
Warr, P. B., 591
Webb, C. H., 501
Weiss, D. J., 543
Wood, L. E., 268
Woodard, W. T., 549

Subjects

(References in *italic* figures are to reviews.)

Abnormal psychology, foundations of, *295–296*
Acoustic interference and 'coding,' 473–476
Acoustic patterns, coding language for, 307–349
Action, *456–457*
Adaptation
—sensory: similarity to salivary habituation, 501–511
—visual: to uniocular image magnification, 66–74
Affective response to sugars, 387–405
Anchoring effects in scaling of value statements, 407–419
Animal: Japanese carp, respiratory conditioning in, 549–554
Animal behavior versus human behavior, *593–595*
Attention: information-processing approach to, *140*; and listening, *452*; and memory, *140;* and performance, *453–454*
Auditory pitch: as a medium, 147–166; scaling of, 147–166
Autokinesis: and egocentric localization, 210–211, 214–217; and experienced displacement, 210–217; and eye fatigue, 211–217; and sensoritonic field theory, 216

Brain and behavior, *455–456*
Brightness, surface: and illumination cues, 1–11

Brightness constancy and illumination apprehension, 1–11

Cerebral dominance, *455*
Channel capacity for bivariate versus univariate stimuli, 205–209
Coding: and acoustic interference, 473–476; and semantic interference, 473–476
Coding language for auditory and visual patterns, 307–349
Cognition, 141–142; development of, in Piaget's theory, *142–144;* models of, *140–142*
Color: psychological effects of, *294;* anthropological study of terms for, *294*
Concept attainment: irrelevant attributes as distractors in, 433; attribute-response correlations in, 425–436; all-or-none versus incremental learning in, 434–435
Concept identification: feedback effects on, 105–106; hypothesis-sampling behavior in, 100–104, 108–111; instructional effects on, 100–111; mediational theory of, 100, 108–111; and problem shifts, 100–111; reversal shifts versus extradimensional shifts in, 100–111
Conceptual processes, memory, and learning, *292–293*
Conditioning
—avoidance: and conditioned-stimulus complexity, 555–563; and intertrial responding, 555–563; in learner versus nonlearner rats, 559–561
—classical: respiratory, in fish, and conditioned-stimulus intensity, 549–554
—operant: effect of pretraining on timing behavior in, 134–139; effects of manifest anxiety on timing behavior in, 134–139
—verbal: *595*
Constancy, brightness: and illumination apprehension, 1–11
Contrast cues versus shadows in surface-lightness perception, 6–8
Critical-trial analysis of recognition learning, 74–84

Development, child: theories of, 142–144; and perceptual learning, 288–292
Displacement, experienced: and autokinesis, 210–217

Effort, muscular: scaling of, 52–65
Errata, 287
Extradimensional shifts: and reversal shifts in concept identification, 100–111

Fatigue, effect on handgrip scaling, 57–65
Form: whole versus part perception of, and pattern complexity, 235–245; orientation of, judgments by children, 218–234

Habituation, salivary: similarities to sensory adaptation, 501–511
Hypothesis sampling in concept identification, 100–104

Illusion: Ebbinghaus', binocularly viewed, 565–566; Müller-Lyer, as two separate illusions, 477–486; size-contrast, without physical size differences, 565–566
inhibition, retroactive: in associative matching, 185–186, 189–193; in free recall, 185–193; in paired-associate learning, 123–133. *See also* Interference
Information: visual processing of, with brief exposure, 355–364; short-term storage of, bivariate versus univariate stimuli, 205–209
Intellect: origins of, *142–144;* Piaget's theory of, *142–144*
Intelligence, artificial, *140–141*
Interference: acoustic, and coding 473–476; in paired-associate learning, 167–172; specific-pair, in free

recall, 185–193; theory of, and serial recall, 112–115; unit-sequence, effects on serial recall, 112–122. *See also* Inhibition

Judgment: of area of rectangles, model for, 543–548; of surface brightness, cues for, 1–11; of form orientation by children, 218–234; of size of diameter and perimeter in contour circles, 437–438

Language, coding: for visual and auditory patterns, 307–349
Learning. *See also* Memory
—applied: *298*
—associative: imagery in, 246–252; and retroactive inhibition, 185–186, 189–193; specific-pair interference in, 185–186, 189–193
—and conceptual processes: *292–293*
—discontinuous: 75–84
—discrimination: visual, dimensional shifts in, 581–586
—meaningfulness and semantic organization in: 253–267
—and memory: molecular theories of, *293–294*
—paired-associate: with the A-B, A-Br paradigm, recovery, retroaction, spontaneous recovery in, 123–133; distributed practice and interstimulus similarity in, 167–172; encoding variability in, 461–471; interference in, 167–172; intralist intrusions in, 276–287; intralist response similarity in, 521–527; intrusions in, 468–469; availability of stimulus and response components of, 12–22; strategies in, 283–286; transfer effects in, 123–133; and interspersed-pair difficulty, 276–287; and meaningfulness, 461–471; and paradigm difficulty, 276–287
—perceptual: and development, *288–292*
—probability: and event probabilities, 180–184; event tracking in, 85–99; subjective estimates of *pi* in, 180–184
—free-recall: associative imagery in, 246–252; specific-pair interference in, 185–193; retroactive inhibition in, 185–193
—serial-order: and interference theory, 112–115; and item frequency, 112–122; unit-sequence interference in, 112–122; versus same-order, rehearsal strategies in, 268–275
—recognition: critical-trial analysis of, 75–84; discontinuities in, 75–84; error computations in, 77–80; error-run reductions in, 80–84
—theories of: *292–293*
—trial-and-error: versus errorless learning, 377–386; effects of stress on, 382–386; training effects in, 382–386; transfer effects in, 382–386
—verbal: *293*, 442–444
Learning task, instrumental, for humans, 567–579
Localization, egocentric: by non-visual method, 214–217; visual, *452*

Meaningfulness: and semantic organization in learning, 253–267; of stimulus and response in paired-associate learning, 461–471
Mechanism and humanism, *296–298*
Meetings: Arrowhead Conference, 442–444; Society of Experimental Psychologists, 439; Southern Society for Philosophy and Psychology, 440–442
Memory: and attention, *140;* and conceptual processes, *292–293;* construction and reconstruction in, 513–519; imagery in, 246–252; immediate, recall-order cuing with different presentation rates, 268–275; information-processing approach to, *140;* long-term versus short-term, coding in, 473–476;

molecular approaches to, 293–294; recognition, for photographs of faces, 351–354; and the semantic organization of related sentences, 253–267; span of, for bivariate versus univariate stimuli, 205–209; and transfer, 293–294; and verbal learning, 293. *See also* Learning
Method: determination of chance alternation levels in the forced-trial paradigm, 421–424; language for pattern coding, 307–349
Movement
—autokinetic: and egocentric localization, 210–211; and experienced displacement with luminous targets, 210–217; and eye fatigue, 211–217; sensoritonic field theory of, 216
—egocentric: 214–217
—patterns of: serial-order errors in, 41
Muscular contraction: phasic, scaling of, 57–65; sustained, scaling of, 52–57, 63–64

Olfaction, 295
Orientation, form: judgments of, by children, 218–234

Patterns: auditory, perceptual coding language for, 307–349; complexity of, in form perception, 235–245; visual, perceptual coding language for, 307–349; movement, serial-order errors in, 41–42
Perception, 456–457; of angle, by visual and haptic means in the matching task, 487–499; of surface brightness, with illumination cues, 1–11; Craik on, *447–449;* of form, whole versus part, and pattern complexity, 235–245; interpersonal, *591–593;* mechanisms of, *589–591;* Piaget on, *589–591;* selective, *452–454;* and knowledge of stimulus alternatives, 173–179; and number of stimulus alternatives, 173–179; spatial localization, *452;* visual, theories of, *449–450*
Perceptual tuning and number of stimulus alternatives, 173–179
Perceptrons, *445–447*
Performance and attention, *453–454*
Personality, theories of, *145*
Phenomenology, Husserl's, *595–596*
Pitch: as a medium, 147–166; scaling of, by transposition, 147–166
Pleasantness of sugars, scaling of, 387–405
Practice, distributed: and interstimulus similarity in paired-associate learning, 167–172
Preference strength: and expected regret ratio in chance situations, 194–204; and expected value difference in chance situations, 194–204; for alternatives in chance situations, 194–204
Problem shifts in concept identification, 100–111
Psychology, Craik on, *447–449*

Response chains, serial-order errors in, 35–49
Response learning and stimulus learning with paired associates, 12–22
Reversal shifts and extradimensional shifts in concept identification, 100–111

Salivation, habituation of: and sensory adaptation, 501–511; and stimulus concentration, 505
Scales, category: of values, 23–34; compared with ratio estimations of values, 23–34
Scaling: anchoring effects in, 407–419; of area of rectangles, model for, 543–548; of diameter and perimeter of contour circles, 437–438; of handgrip contraction, phasic and sustained, 52–65; of pitch, by musical and nonmusical subjects, 149–166; of pitch, by transposition methods, 147–166; of pleasantness of sugars, 387–

405; of sweetness of sugars, 387–405; of values, anchoring effects in, 407–419
Selective perception, *452–454*
Semantic organization and memory for related sentences, 253–267
Sensoritonic field theory and autokinesis, 216
Serial-order errors in response chains, 35–49
Shadows, surface lightness, and perception, 1–6, 9–10
Shifts: dimensional, in visual-discrimination learning, 581–586; extradimensional, in concept identification, 100–111; problem, in concept identification, 100–111; reversal, and extradimensional shifts in concept identification, 100–111
Sleep, *455–456*
Speech behavior and stress pre-entry, 35–41
Split-brain research, *454–455*
Spontaneous recovery: in A-B, A-Br paired-associate paradigm, 123–137; and stimulus meaningfulness in paired-associate learning, 123–133
Spoonerisms: as stress pre-entry, 35–41; experimentally produced, 38–41
Stress, effects on trial-and-error versus errorless learning, 382–386
Stress pre-entry: and finger-movement patterns, 41–44; model of, 44–47; and spoonerisms in speech behavior, 35–41; theories of, 44–49
Surface lightness: and cues for illumination, 1–11; perception of, 1–11
Sweetness of sugars: and chemical structure, 398–401; scaling of, 387–405

Taste: sweetness of sugars, scaling of, 387–405

Values, scaling of: anchoring effects in, 407–419; by category ratings, 23–34; by ratio estimates, 23–34
Vigilance behavior: and display coupling, 529–532, 536–540; versus matching-session behavior, 533–536; modality effects on, 529–540; and task difficulty, 529–530, 539–540
Vision: fundamentals of, *450–452;* theories of, *449–450*
Visual adaptation and uniocular image magnification, 66–74
Visual-discrimination learning, dimensional shifts in, 581–586
Visual form perception of angles, 487–498
Visual hemifield effects with brief exposures and hemifield reports, 355–364
Visual patterns, coding language for, 307–349
Visual persistance of a moving object under stroboscopic illumination, 365–375
Visual scanning with brief exposures, 355–364
Von Restorff hypothesis of serial-order errors, 43–44

Vol. 84 Founded in 1887 by G. Stanley Hall 1971

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys
Charles W. Eriksen

M. E. Bitterman
Benton J. Underwood

Volume 84

1971

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS
Urbana, Illinois 61801

TABLE OF CONTENTS

Articles and Notes

Anderson, N. H., and Weiss, D. J. Test of a Multiplying Model for Estimated Area of Rectangles, 543

Appelle, S. Visual and Haptic Angle Perception in the Matching Task, 487

Arthur, B. *See* Coltheart, M.

Attneave, F., and Olson, R. K. Pitch as a Medium: A New Approach to Psychophysical Scaling, 147

Beck, J. Surface Lightness and Cues for the Illumination, 1

Binford, J. R. *See* Loeb, M.

Bishop, G., and Smith, J. Dimensional Shifts with Verbal Units, 581

Bolt, M. Imagery in Associative Learning and Memory, 246

Cain, W. S., and Stevens, J. C. Effort in Sustained and Phasic Handgrip Contractions, 52

Carey, S. T. *See* Martin, E.

Cohen, J. *See* Lorei, T. W.

Coltheart, M., and Arthur, B. Tachistoscopic Hemifield Effects with Hemifield Report, 355

Coren, S. A Size-Contrast Illusion without Physical Size Differences, 565

Damianopoulas, E. N. Stimulus and Response Availability in Associative Learning, 12

Delprato, D. J. Specific-Pair Interference on Recall and Associative-Matching Retention Tests, 185

Dey, M. K. Distributed Practice in Paired-Associate Learning: Interstimulus Similarity, 167

Donnelly, M. B. *See* Rimoldi, H. J. A.

Efton, R., and Lee D. N. The Visual Persistence of a Moving Stroboscopically Illuminated Object, 365

Epstein, W. Adaptation to Uniocular Image Magnification after Varying Preadaptation Activities, 66

Erikson, J. R. Problem Shifts and Hypothesis Behavior in Concept Identification, 100

Erlebacher, A. *See* Sekuler, R.

Fischer, G. J. Sum of Event Probabilities and Subjective Estimates of *Pi* in Probability Learning, 180

Fleming, J. S. *See* Summers, S. A.

Galper, R. E., and Hochbert, J. Recognition Memory for Photographs of Faces, 351

Glick, J. Is Autokinetic Motion Related to Experienced Displacement? 210

Gottwald, R. L. Attribute-Response Correlations in Concept Attainment, 425

Gruneberg, M. M., and Sykes, R. N. 'Coding' in Studies of Acoustic and Semantic Interference and Confusion, 473
Gummerman, K. Selective Perception and the Number of Alternatives, 173
Hall, J. F. Formal Intralist Response Similarity: Its Role in Paired-Associate Learning, 521
Harris, L., and Schaller, M. J. Form and its Orientation: Reexamination of A Child's-Eye View, 218
Hinrichs, J. V. *See* Wood, L. E.
Hochberg, J. *See* Galper, R. E.
Hui, C. *See* Liu, I.
Johnson, G. J. Transfer as A Function of Paradigm and Interspersed-Pair Difficulty, 276
Laack, W. J. *See* Rimoldi, H. J. A.
Lee, D. N. *See* Efron, R.
Lee, W. The Effects of Expected Value Difference and Expected Regret Ratio on Preference Strength, 194
Leeuwenberg, E. L. J. A Perceptual Coding Language for Visual and Auditory Patterns, 307
Levis, D. J. CS Complexity and Intertrial Responding in Shuttlebox Avoidance Conditioning, 555
Liu, I., and Hui, C. An Instrumental-Learning Task for Human Subjects, 567
Loeb, M., and Binford, J. R. Modality, Difficulty, and 'Coupling' in Vigilance Behavior, 529
Lorei, T. W., and Cohen, J. A Comparison of Two Methods of Value Measurement, 23
McBurney, D. H. *See* Webb, C. H.
MacKay, D. G. Stress Pre-Entry in Motor Systems, 35
Manning, S. K. A Method to Assess Chance Alternation Levels in the Forced-Trial Paradigm, 421
Martin, E., and Carey, S. T. Retroaction, Recovery, and Stimulus Meaningfulness in the A-B, A-Br Paradigm, 123
Millward, R. B. *See* Reber, A. S.
Moskowitz, H. R. The Sweetness and Pleasantness of Sugars, 387
Olson, R. K. *See* Attneave, F.
Pasnak, R. Pattern Complexity and Response to Distinctive Features, 235
Postman, L., and Stark, K. Encoding Variability and Transfer, 46
Prather, D. C. Trial-and-Error versus Errorless Learning: Training, Transfer, and Stress, 377
Reber, A. S., and Millward, R. B. Event Tracking in Probability Learning, 85
Rimoldi, H. J. A., Donnelly, M. B., and Laack, W. J. Anchoring Effects in Psychological Scaling, 407
Sasson, R. Y. Semantic Organizations and Memory for Related Sentences, 253
Schaller, M. J. *See* Harris, L.
Sekuler, R., and Erlebacher, A. The Two Illusions of Müller-Lyer: Confusion Theory Reexamined, 477
Singh, D. Effect of Level of Manifest Anxiety and Type of Pretraining on Timing Behavior, 134
Slak, S. Univariate and Bivariate Information Storage, 205
Smith, J. *See* Bishop, G.

Stark, K. *See* Postman, L.
Steinmetz, J. I. *See* Turnage, T. W.
Stevens, J. C. *See* Cain, W. S.
Summers, S. A., and Fleming, J. S. Construction and Reconstruction in Memory, 513
Sykes, R. N. *See* Gruneberg, M. M.
Teghtsoonian, R., and Teghtsoonian, M. Discontinuities in Recognition Learning Revealed by Critical-Trial Analysis, 75
Teghtsoonian, R., and Teghtsoonian, M. The Apparent Lengths of Perimeters and Diameters Define a Ratio Smaller than Pi, 437
Turnage, T. W., and Steinmetz, J. I. Unit-Sequence Interference and Short-Term Recall, 112
Webb, C. H., and McBurney, D. H. Salivary Habituation: Quantitative Similarities to Sensory Adaptation, 501
Weiss, D. J. *See* Anderson, N. H.
Wood, L. E., and Hinrichs, J. V. Cuing of Recall Order in Immediate Memory with Different Rates of Presentation, 268
Woodard, W. T. Classical Respiratory Conditioning in the Fish: CS Intensity, 549

Book Reviews

(The reviewer's name appears in parentheses after the title of the work.)

Andreas, B. G. *Psychological Science and the Educational Enterprise* (E. Z. Rothkopf), 298
Berlin, B., and Kay, P. *Basic Color Terms: Their Universality and Evolution* (R. W. Burnham), 294
Bertalanffy, L. von *Robots, Men and Minds* (T. G. Bever), 296
Birren, F. *Light, Color, and Environment* (R. W. Burnham), 295
Byrne, W., ed. *Molecular Approaches to Learning and Memory* (J. A. Deutsch), 293
Craik, K. J. W., ed. Sherwood, S. L. *The Nature of Psychology: A Selection of Papers, Essays, and Other Writings* (H. F. Crovitz), 447
Das, J. P. *Verbal Conditioning and Behavior* (R. F. Hefferline), 595
Elkind, D., and Flavell, J. H., eds. *Studies in Cognitive Development: Essays in Honor of Jean Piaget* (D. Kuhn), 142
Elyeton, R. O., ed. *The Phenomenology of Husserl: Selected Critical Readings* (W. Cohen), 595
Flavell, J. H. *See* Elkind, D.
Furth, H. G. *Piaget and Knowledge* (D. Kuhn), 142
Gazzaniga, M. S. *The Bisected Brain* (C. Blakemore), 454
George, F. H. *Models of Thinking* (H. F. Crovitz), 140
Gibson, E. J. *Principles of Perceptual Learning and Development* (A. Gilinsky), 288

Hazlitt, H. *Thinking as a Science* (H. F. Crovitz), 141
Jarvinen, J. *Orthogonal Localization of Visual Objects: Some Experiments on Space Perception* (W. B. Templeton), 452
Kay, P. *See* Berlin, B.
Keppel, G. *See* Postman, L.
Kintsch, W. *Learning, Memory, and Conceptual Processes* (J. Deese), 292
Knapper, C. *See* Warr, P. B.
London, P., and Rosenhan, D., eds. *Foundations of Abnormal Psychology* (H. J. Johnson), 295
Lorenz, K. *Studies in Animal and Human Behavior*. Vol. I (E. Klinghammer), 593
McCartney, W. *Olfaction and Odours: An Osphrésiological Essay* (R. J. O'Connell), 295
Maier, H. W. *Three Theories of Child Development* (D. Kuhn), 142
Mehrabian, A. *An Analysis of Personality Theories* (R. D. Singer), 145
Minsky, M., and Papert, S. *Perceptrons* (E. Hunt), 445
Moray, N. *Listening and Attention* (U. Neisser), 452
Norman, D. A. *Memory and Attention* (H. Egeth), 140
Papert, S. *See* Minsky, M.
Pastore, N. *Selective History of Theories of Visual Perception, 1650–1950* (S. Coren), 449
Phillips, J. L., Jr. *The Origins of Intellect: Piaget's Theory* (D. Kuhn), 142
Piaget, J. *The Mechanism of Perception* (S. Coren), 589
Postman, L., and Keppel, G., eds. *Verbal Learning and Memory: Selective Readings* (J. Deese), 293
Pribram, K. H., ed. *Brain and Behavior*. Vol. 1, *Mood, States, and Mind* (J. R. Averill), 455
Pribram, K. H., ed. *Brain and Behavior*. Vol. 2, *Perception and Action* (I. P. Howard), 456
Rosenhan, D. *See* London, P.
Rubin, M. L., and Walls, G. L. *Fundamentals of Visual Science* (D. C. Hood), 450
Sanders, A. F., ed. *Attention and Performance III* (D. I. Mostofsky), 453
Sherwood, S. L., ed. *See* Craik, K. J. W.
Walls, G. L. *See* Rubin, M. L.
Warr, P. B., and Knapper, C. *The Perception of People and Events* (P. Shaver), 591

EDUCATIONAL AND PSYCHOLOGICAL
LIBRARY

"In a moment you will fall deeply asleep. When you do, a dream will form. You will find yourself sitting in a movie theater looking at a blank screen. . . ."

With these words and careful suggestions from the therapist, the patient will fall into a hypnotic sleep and will redream the previous night's experience, in a form of psychotherapy examined in

DREAMS, IMAGES, AND FANTASY

A Semantic Differential Casebook

by C. Scott Moss, with a foreword by Charles E. Osgood

"Moss writes with considerable expertise on the subject of how hypnosis can effectively be used to facilitate psychotherapy that is intensive, time limited, and crisis oriented. He artfully combines the understanding gained from a highly regarded quantitative measurement of meaning, Osgood's semantic differential with his own clinical sensitivities, to illustrate the value of hypnosis in reaching and utilizing the symbol-translating mechanisms of the mind. This book is unique in the field. . . ." — *Choice*. $8.50

"BLACK ROVER, COME OVER!"

The Hypnosymbolic Treatment of a Phobia

by C. Scott Moss and Pauline James

Published to accompany *Dreams, Images, and Fantasy,* this book stands on its own as an account of the swift and successful treatment of a patient with an extreme phobia. It "consists largely of James's personal account of her therapeutic experience, accompanied by brief notes by psychologist Moss of his interpretations of the same therapy sessions. . . . a fascinating, well written story." — *Library Journal*. ". . . a fascinating view of what may be accomplished in the alleviation of symptoms and the development of personal freedom when a brilliant, dedicated therapist cautiously utilizes hypnotic techniques to elucidate dream material." — *American Journal of Orthopsychiatry*. Paper, $3.95. With audiotape of a transcribed therapy session (1 hour, 45 minutes, 3¾ IPS, open reel or cassette), $12.50

UNIVERSITY OF ILLINOIS PRESS up **Urbana Chicago London**

CRAIG R. EISENDRATH

THE UNIFYING MOMENT

The Psychological Philosophy of William James and Alfred North Whitehead

Mr. Eisendrath merges James's theory of will and personality into Whitehead's theory of concresence and organism to achieve a result which is both an extension of James and an exegesis of Whitehead. James's psychology is grounded in philosophic doctrine, and Whitehead's abstract philosophy given greater specificity through the concrete illustrations of psychology. $10.00

In Canada/Saunders of Toronto, Ltd.

HARVARD
UNIVERSITY PRESS

REHABILITATION INDIA
CALCUTTA
700[illegible]